ADA 123753

DEVELOPMENT AND PRODUCTION COST ESTIMATING RELATIONSHIPS FOR AIRCRAFT TURBINE ENGINES

J. L. Birkler, J. B. Garfinkle, K. E. Marks

October 1982

N-1882-AF

The United States Air Force

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER N-1882-AF	2. GOVT ACCESSION NO. AD-A123 753	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Development and Production Cost Estimating Relationships for Aircraft Turbine Engines		5. TYPE OF REPORT & PERIOD COVERED Interim 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) J. L. Birkler, J. B. Garfinkle, K. E. Marks		8. CONTRACT OR GRANT NUMBER(s) F49620-82-C-0018
9. PERFORMING ORGANIZATION NAME AND ADDRESS The Rand Corporation 1700 Main Street Santa Monica, CA. 90406		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Requirements, Programs & Studies Group (AF/RDQM) Ofc, DCS/R&D and Acquisition Hq USAF, Washington, D.C. 20330		12. REPORT DATE October 1982 13. NUMBER OF PAGES 68
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release: Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

No Restrictions

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Cost Estimates	Engines
Expenses	Turbines
Aircraft Engines	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

See reverse side

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

This Note describes Rand's latest study of cost estimating relationships for new military aircraft turbine engine development and production programs. It presents equations for estimating development and production costs and time of arrival for U.S. military turbojet and turbofan engines. The study ~~dervies~~ new cost estimating relationships from an expanded data base and uses new diagnostic statistics to screen the relationships and to evaluate the characteristics of the preferred set. Section ~~II~~ of this Note identifies the data used, explains the criteria and rationale for selecting explanatory variables, and describes recently developed regression diagnostics. Section III presents the preferred set of relationships. Comments on these results; a comparison with DAPCA equations; suggestions for the use of the cost estimating relationships and directions for possible future research are discussed in Section IV. Supporting statistics for the predictive models are available in the Appendix.

A RAND NOTE

DEVELOPMENT AND PRODUCTION COST ESTIMATING RELATIONSHIPS FOR AIRCRAFT TURBINE ENGINES

J. L. Birkler, J. B. Garfinkle, K. E. Marks

October 1982

N-1882-AF

Prepared for The United States Air Force

Rand
SANTA MONICA, CA. 90406

PREFACE

This Note describes Rand's latest study of cost estimating relationships for new military aircraft turbine engine development and production programs. The recent availability of data for several new engine programs and new, more powerful statistical tools motivated us to develop turbine engine cost estimating methods that are simple and easy to use, but accurate enough for long range planners. Such methods are suitable for conceptual exercises, planning studies, Independent Cost Analyses (ICA), and other situations for which conventional detailed estimating procedures are either impractical or overly time-consuming. The estimating relationships developed should be of interest to those persons throughout the Air Force and elsewhere in DOD who are concerned with long range planning and the preparation or review of turbine engine development and production program costs.

This research was originally undertaken as part of the Project AIR FORCE project "Cost Analysis Methods for Air Force Systems," which has since been superseded by "Air Force Resource and Financial Management Issues for the 1980s" in the Resource Management Program.

SUMMARY

This Note presents equations for estimating development and production costs and time of arrival for U.S. military turbojet and turbofan engines. Interest in further investigation of aircraft turbine engine cost estimating relationships (CERs) grew out of the availability of data for engines recently developed, and experience with the CERs in Rand's computer model for estimating Development and Procurement Costs of Aircraft (DAPCA).

After establishing criteria for selecting explanatory variables and CERs, regression analysis was applied to the expanded data base to develop improved relationships for the cost of development to the model qualification test (MQT), total development cost, and the cumulative average price at the 1000th production engine. The engine characteristics that best explain development cost through MQT and production cost are maximum thrust of the engine at sea-level-static conditions, an indicator of engine size; Mach number, a measure of performance; and turbine inlet temperature, the dominant technical parameter in the engine cycle. For total development cost, which includes the expenses involved in developing a new engine to MQT, plus the cost to correct service related deficiencies and costs for continual performance and reliability improvements over time, the derived equation includes a production quantity term as well as thrust and Mach number.

The estimating relationship for time of arrival (TOA) was also refined in this study. (The TOA method links certain engine performance characteristics with time to provide a measure of an engine's state of

the art). The refined TOA model is based on 29 U.S. military turbojet and turbofan engines developed and produced during the past 30 years. The model predicts the man-rated MQT date as a function of certain of the engine's performance and design parameters. The parameters include engine thrust to weight ratio, turbine inlet temperature, and specific fuel consumption, which are the three most important technical characteristics in the turbine engine development process.

Several new diagnostic statistics, generated as part of the analysis, provide more insight into the estimation error and into the influence of specific data points on the derived equations than was available in earlier studies. These diagnostics should help estimators understand the sensitivities of the CERs, and therefore the estimated costs, to particular characteristics of individual engines in the data base. Coupled with the expanded data base and selection criteria, these regression diagnostics have also helped identify CERs that intuitively satisfy our engineering sense and generally have fewer explanatory variables, while improving their predictive capability.

These models are intended for use by long range military planners attempting to determine costs for new systems--especially those of a technically advanced nature--so that better estimates can be made. All parameters needed are readily available at an early stage of planning for a new system. Care must be exercised in using these models to ensure that inputs are consistent with the data base used in this study. For example, cost estimates will reflect military technology and the manner in which programs were conducted during the 1950s, 1960s, and 1970s. If an engine is developed that is not in the mainstream trend,

such as a variable cycle or lift engine, the estimating relationship described may not apply. To the extent that a new program differs from historical conditions, extrapolation will be necessary.

ACKNOWLEDGMENTS

Personnel in several corporations and military organizations contributed the turbine engine technical and cost data that are the basis for this research. Most of the data were originally provided to Rand for studies conducted earlier, but data for new engines were provided specifically for this research by the Aeronautical Systems Division, Air Force Systems Command, and the Naval Air Systems Command. These new data are the key to much of the improved estimating capability provided by the models derived in this study. R. Thompson and R. Runkel, Headquarters Aeronautical Systems Division; Major M. Shutak, Headquarters Air Force Accounting and Finance Center, along with A. Pressman, Naval Air Systems Command, gave generously of their time and made substantitive comments while reviewing an earlier draft. Rand colleagues, J. R. Nelson, H. G. Massey, and W. Rogers provided constructive comments that have led to many improvements.

SYMBOLS

Symbol	Definition
AIR	Airflow through the engine in lbs/sec at maximum rated thrust.
CIP	Component Improvement Program
DEVTIME	Development time from start to MQT, months
FFER	Full Flight Envelope Release
IFR	Initial Flight Release
ISR	Initial Service Release
MACH	Maximum flight envelope Mach number (measure of speed related to speed of sound); 1.0 for engines designed for subsonic flight.
MQT	Model Qualification Test
MQTDEVCOST	Development cost to MQT, millions of FY 1980 dollars
OCR	Operational Capability Release
PFRT	Preliminary Flight Rating Test
PROCOST	Cumulative average production cost through the 1000th engine; thousands of FY 1980 dollars.
QMAX	Maximum dynamic pressure in flight envelope, lb/ft^2
QTY	Quantity of engines produced.
SFCMIL	Specific fuel consumption at military thrust, sea-level static conditions (lb/hr/lb thrust).
TEMP	Maximum turbine inlet temperature (degrees Rankine).
THRAIR	Thrust to air flow ratio (THRMAX/AIR).
THRMAX	Maximum rated thrust at sea-level static conditions, including afterburner thrust if any (lb).
THRWGT	Thrust to weight ratio (THRMAX/WGT).

TOA	Time of arrival at successful MQT (in calendar quarters since the third quarter of 1942).
TOTDEVCOST	Total cost of development (millions of FY 1980 dollars) including development to MQT and product improvement, evaluated at dates of MQTs other than the first for a particular engine model.
TOTPRS	Engine pressure term (psf) computed as the product of engine maximum pressure ratio and the maximum dynamic pressure of the engine design envelope.
WGT	Engine dry weight (lb).
Δ TOA	The interval between TOA and the actual date an engine passes its MQT. (Δ TOA = calculated TOA - actual TOA date).

CONTENTS

I. INTRODUCTION

A modern military aircraft consists of four main subsystems: airframe, avionics, propulsion and armament. The propulsion system of most new military aircraft is a turbine engine. Modern turbine engines are complex and costly. A turbine engine can cost over a billion dollars for development and product improvement over its lifetime and accounts as much as 25 percent of the flyaway cost of a fighter aircraft. It follows that reasonably accurate estimates of engine development and production cost are needed during planning for a new engine in an aircraft system.

For conceptual planning studies, preliminary tradeoff analyses, and the like, it would be desirable to have a simple and easy to use procedure for estimating the cost of conceptual or proposed turbine engines within ± 25 percent. Such a procedure should allow the analyst to obtain estimates at a time when descriptive information (e.g., engine characteristics) is imperfect and often incomplete, and there are no engine development program data (e.g., number of test engines and test hours).

Consequently Rand has, over the years, conducted a number of studies of turbine engine cost estimation. One of the most widely known efforts produced cost estimating relationships (CERs) that were later incorporated into a computer model of aircraft system costs generally known as DAPCA (Development and Procurement Costs of Aircraft).[1]

---

[1] J. R. Nelson and F. S. Timson, Relating Technology to Acquisition Costs: Aircraft Turbine Engines, R-1288-PR, The Rand Corporation, March 1974. The estimating relationships developed in this research are incorporated in the DAPCA model: H. E. Boren, Jr., A Computer Model for Estimating Development and Procurement Costs of Aircraft (DAPCA III), R-1854-PR, The Rand Corporation, March 1976.

Several circumstances make it appropriate now to take another look at engine development and production costs. Experience with the DAPCA engine equations has shown them to be quite sensitive to small changes or errors in the parameters used to drive the estimates. When using DAPCA, some investigators have also reported a tendency toward underestimation for the latest high performance engines. Furthermore, three new engines have been developed since the DAPCA equations were derived. Adding these engines to the analysis offered two advantages: First, we can expect that they are more similiar to the next generation of turbine engines that will be developed, and second, increasing the sample size provides additional confidence in the predictive capability of the CERs. Recently, some regression diagnostics have become available that greatly increase the amount of information generated during regression analysis. Several of these provide useful information about collinearity and the influence of individual data observations on the various estimated parameters. These statistics seemed well suited to evaluating the difficulties some analysts were encountering with the DAPCA equations.

The hypothesis of this study was that using recent engine data along with newly developed statistical methods would make it possible to develop more stable and accurate CERs that would be useful in predicting costs of large, modern turbofan and turbojet aircraft engines. Such equations would be used only before or early in the planning and development phase of an engine program, when no drawings or bills of material exist. By the time development is well along and higher resolution estimates are needed, greater accuracy can be achieved by

extrapolating from actual costs of nameplate engines than the parametric methods described here.

Historically, the engine program milestones distinguished for estimating purposes were as follows:

- <u>Preliminary Flight Rating Test (PFRT)</u>. A series of individual tests that in combination demonstrated that the engine was suitable for use in experimental flight testing.

- <u>Model Qualification Test</u> (<u>MQT</u>). A series of individual tests that in combination demonstrated that the engine was suitable for production.

- <u>Delivery of the Nth Engine</u>. The time when some number, say 1000, of production engines have been delivered.

Difficulties encountered with recent aircraft turbine engines have led to some important changes in development emphasis and new test procedures to insure delivery of more supportable engines. The Air Force has shifted to a four-step development process. The proposed milestones in this process are:

- <u>Initial Flight Release</u> (<u>IFR</u>). A series of individual tests that in combination demonstrate the engine is suitable for limited flight testing.

- <u>Full Flight Envelope Release</u> (<u>FFER</u>). A series of individual tests that in combination demonstrate that the engine is suitable for flight testing throughout the full aircraft performance flight envelope.

- <u>Initial Service Release</u> (<u>ISR</u>). A series of individual tests that in combination demonstrate the engine is suitable for low rate production.

- <u>Operational Capability Release</u> (<u>OCR</u>). A series of individual tests that in combination demonstrate the engine is suitable for full production release.

No engines have yet been developed under the new four-step development process. This fact coupled with the availability of historical data in traditional formats led us to continue using traditional engine development milestones. Nevertheless, we can offer some suggestions on making the transition between the two approaches.

The old PFRT milestone would roughly correspond to IFR or step one. The MQT milestone becomes harder to identify, but should approximate ISR or step three. Step four historically would have occurred in the early stages of the Component Improvement Program (CIP). However, under the PFRT/MQT development concept, additional requirements (and thus costs) were continually added as the military specifications evolved over the years. It can be considered that the four-step development process merely represents further evolution and that any additional costs incurred between ISR and OCR will be captured within the accuracy error of the model.

This study derives new CERs from an expanded data base and uses new diagnostic statistics to screen the CERs and to evaluate the characteristics of the preferred set. Section II of this Note identifies the data used, explains the criteria and rationale for selecting explanatory variables, and describes recently developed regression diagnostics. Section III presents the preferred set of CERs. Comments on these results; a comparison with DAPCA equations; suggestions for the use of these CERs and directions for possible future research are discussed in Sec. IV. Supporting statistics for the predictive models are available in the appendix.

II. RESEARCH PROCEDURE

Before we review the data base and the analytical procedure used, it is appropriate to consider the objectives of this analysis and the limitations they impose. Engine development and production costs are ordinarily estimated by manufacturers in great detail, frequently by the application of standard engineering and manufacturing hours to each operation and building up to a total. Whatever the degree of accuracy obtained through this process, it is very time-consuming and requires a detailed knowledge of the development and manufacturing process as well as up-to-date information on standard hours and material requirements and costs for the various fabrication, assembly, and test procedures. For planning studies, preliminary tradeoff analysis, and the like, it is desirable to have a simple procedure for estimating engine costs at important program milestones.

Early in the planning cycle, when resources are limited and detailed knowledge of design specification is unavailable, parametric estimating models requiring few inputs have been found to provide cost estimates that are sufficiently accurate for these purposes. Parametric estimating relationships were sought in this study for four dependent variables: the cost of development up to successful completion of MQT (MQTDEVCOST); total development cost (TOTDEVCOST), which includes MQT and all subsequent product improvement during the life of the engine; 1000th unit cumulative average production cost (PROCOST); and time of arrival (TOA) at MQT. The last term attempts to quantify the level of technology for a given engine development program. Unlike some earlier

Rand studies, time of arrival is not used as an input to the cost estimating equations derived here. Rather it is provided to give the estimator an indication of the degree of risk associated with an engine program.

TECHNICAL DATA

Technical data came from several sources. Most were originally provided to Rand for studies conducted earlier, but other data were provided directly from the various military offices and manufacturers involved in the development and production of individual engines. Table 1 presents the data used to develop the MQTDEVCOST, PROCOST, and TOA relationships.[1] Three engines have been added to the data base--the F100, F101, F404.

The engine characteristics used are those of a particular model, depending on the context. Aircraft engines evolve through a succession of different versions. Normally, several versions are in production at the same time and on the same production line. When development costs are examined through MQT, the engine characteristics used are those of the first production model. When production costs are examined, however, often during the course of an engine production program, new models are introduced having performance and technical attributes that differ substantially from the original version. For these cases, a subjective selection of the most representative series engine was made.

Table 2 gives the technical and production quantity data used in the TOTDEVCOST model. Data for the engine series that passed the first MQT for each model are listed first. Subsequent entries for the same

---

[1] In order to make the report suitable for general distribution, proprietary cost and classified technical data are not shown.

Table 1

TECHNICAL DATA FOR U.S. MILITARY AIRCRAFT TURBINE ENGINES

Engine	Turbine Inlet Temp. (^oR) TEMP	Weight (lb) WGT	Pressure Term (lb/ft^2) TOTPRS	Specific Fuel Consumption (lb/hr/lb) SFCMIL	Mach No. MACH	TOA (qtr)	Airflow (lb/sec) AIR	Max. Thrust THRMAX
F100	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
F101	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
F404	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
TF30 (b)	2430	3850	51340	0.63	2.2	92	240	18500
TF30 (c)	2540	4112	51340	0.63	2.5	92	240	20840
TF33	2060	3900	19240	0.52	1.0	71	458	17000
TF34	2660	1420	16500	0.36	1.0	120	338	9275
TF39	2840	7300	19500	0.32	1.0	109	1555	40800
TF41	2620	3175	28770	0.65	1.0	107	258	14500
J30	1830	686	1575	1.17	1.0	17	--	1560
J31	1930	850	1710	1.25	1.0	11	--	1600
J33	1960	1875	3400	1.22	1.0	19	76	3825
J34	1895	1200	3400	1.06	1.0	27	--	3250
J35	2010	2300	3400	1.08	1.0	21	72	4000
J40	1985	3580	5750	1.08	1.8	45	--	10900
J42	1825	1729	3640	1.25	1.0	25	--	5000
J46	1985	1863	6625	1.01	1.8	44	--	6100
J47	2060	2475	5375	1.10	1.0	26	91	4850
J48	2030	2040	4880	1.14	1.0	33	112	6250
J52	2060	2050	12840	0.82	1.8	74	122	8500
J57	2060	4160	11400	0.80	1.4	41	162	10000
J58	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
J60	2060	460	10360	0.96	1.0	71	50	3000
J65	2030	2815	8500	0.92	1.2	46	117	7220
J69	1985	364	3400	1.12	1.0	44	21	1025
J71	2160	4090	11000	0.88	1.5	47	155	9570
J73	2060	3825	8750	0.92	1.9	49	142	8920
J75	2060	5950	16724	0.80	2.0	59	252	23500
J79	2160	3225	18056	0.87	2.0	57	162	15000
J85	2100	570	10360	1.03	2.0	74	42	3850

[a] Data removed for security reasons.

[b] Used in Development Cost CER.

[c] Used in Production Cost CER.

Table 2

U.S. MILITARY TURBINE ENGINE TECHNICAL AND PRODUCTION QUANTITY DATA USED TO DERIVE TOTAL DEVELOPMENT COST EQUATIONS

Engine	Turbine Inlet Temp. (^oR) TEMP	Weight (lb) WGT	Pressure Term (lb/ft^2) TOTPRS	Specific Fuel Consumption (lb/hr/lb) SFCMIL	Mach No. MACH	TOA (qtr)	Airflow (lb/sec) AIR	Max. Thrust THRMAX	DEVTIME (mo)	Engines Produced (QTY)
F100	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)	22	0
F100	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)	22	697
F100	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)	34	1700
TF30	2430	3852	51340	0.63	2.2	92	240	18500	21	0
TF30	2540	4112	52850	0.63	2.5	104	243	20840	33	1000
TF30	2610	4027	52850	0.69	2.5	115	260	25100	44	2500
TF33	2060	3900	19240	0.52	1.0	71	458	17000	9	0
TF33	2210	4605	23680	0.61	1.0	82	498	21000	20	1200
TF33	2210	4605	23680	0.61	1.0	111	498	21000	49	2750
TF34	2660	1420	16500	0.36	1.0	120	338	9275	18	0
TF34	2660	1420	16500	0.36	1.0	140	338	9275	48	651
J52	2060	2050	12840	0.82	1.8	74	122	8500	29	0
J52	2160	2118	19240	0.86	1.8	91	137	9300	46	1700
J52	2460	2318	21608	0.89	1.8	111	143	11200	66	3500
J57	2060	4160	11400	0.80	1.4	41	162	10000	22	0
J57	2060	5160	11400	0.84	1.4	51	172	16000	32	1500
J57	2135	4750	11400	0.83	1.4	59	183	16900	40	6500
J60	2060	460	19360	0.96	1.0	71	50	3000	11	0
J60	2060	460	10360	0.96	1.0	96	50	3300	36	800
J75	2060	5950	16724	0.80	2.0	59	252	23500	20	0
J75	2070	5875	17612	0.82	2.0	67	252	24500	28	400
J75	2070	5875	17612	0.82	2.0	84	252	24500	45	1400
J79	2160	3225	18056	0.87	2.0	57	162	15000	18	0
J79	2160	3375	18056	0.84	2.0	63	162	15800	24	400
J79	2235	3675	24375	0.67	2.0	72	169	17000	33	1150
J79	2270	3850	24375	0.84	2.0	98	170	17900	59	7000
J85	2100	570	10360	1.03	2.0	74	42	3850	26	0
J85	2160	587	10360	1.03	2.0	85	44	4080	37	1800
J85	2200	600	10360	1.04	2.0	98	44	4300	50	3700

[a] Data removed for security reasons.

engine denote a data value corresponding to one of a series of MQTs in the continuing development of that engine.

Tables 1 and 2 reflect the consideration in this study of turbojet and turbofan engines only. Turboprop and turboshaft engines included in some earlier Rand analyses have been excluded. For future Air Force systems, these engines are expected to be of less relevance than turbofan or turbojet engines.

COST DATA

With data collected from diverse sources, comparability of costs becomes a serious problem. After insuring that all definitional differences were eliminated, three major adjustments were made--one for price-level changes, one for quantity, and one for program peculiarities--so that comparisons could be made on the basis of constant dollars, at a specified production quantity, and for generally similar acquisition strategies.

Price-Level Adjustment

The term "cost" as used in this Note refers to the total price to the government expressed in Fiscal Year (FY) 1980 dollars of an engine development or production program. All costs were adjusted to this price level with the index shown in Table 3. This index, made up of several weighted labor and material sub-indices, was developed specifically for aircraft turbine engines by the Air Force Systems Command Cost Analysis Improvement Group (AFSC/CAIG) in 1981.[2] Use of

---

[2] A. Fatkin, AFSC/CAIG Research REPORT NR1 Generic Inflation Indexes for Weapon Systems, Hq, AFSC/ACCE, July 1981.

this index to adjust total engine cost implies that the development process is similar to the production process, which it is not. The development of a specific RDT&E index was beyond the scope of this effort, however. Consequently, we have used a single index as a primary indicator of price movement.

Adjustment for Quantity

Some of the engines in the data sample were produced in thousands and others in hundreds. To compare these engines, it was necessary to establish production cost baselines for all engines at a single production quantity (1000 units). For those engines that were not

Table 3

AFSC/CAIG AIRCRAFT ENGINE COST INDEX

Year	Index	Year	Index
1946	6.44	1964	3.25
1947	6.16	1965	3.17
1948	5.60	1966	3.06
1949	5.42	1967	2.95
1950	5.13	1968	2.82
1951	4.44	1969	2.69
1952	4.19	1970	2.52
1953	4.15	1971	2.39
1954	4.06	1972	2.29
1955	3.88	1973	2.14
1956	3.60	1974	1.97
1957	3.50	1975	1.69
1958	3.53	1976	1.56
1959	3.49	1977	1.42
1960	3.45	1978	1.31
1961	3.40	1979	1.17
1962	3.37	1980	1.00
1963	3.31		

produced in large quantities (or their production runs had not yet reached 1000 units--F404) costs were extrapolated along the established slope or, in the case for the F101 engine, estimated values were obtained from the USAF.

Program Adjustment

Although each engine program is unique, the F100 and F404 had such unusual characteristics that we adjusted these costs to make them comparable to other programs.

The F100 engine development program, for example, began as a joint effort with both the USAF and Navy developing a common engine core that was planned to evolve into the F100 for use in the F-15 and the F401 for use in the F-14B. Total development dollars spent by both services were more than that needed for a single engine, but less than that required to develop two engines. Thus the development cost values used for the F100 engine does not include those dollars spent on the Navy version.

An adjustment was also made to the F404 development cost. The F404 engine, powerplant for the Navy's F-18, had its genesis as the YJ101, a powerplant funded by the Air Force and used in the YF-17. Just using Navy F404 development dollars ignores the fact that considerable Air Force resources were spent on the YJ101 that eventually benefited the F404. In this case, official Navy F404 development costs were adjusted upward. This adjustment was made based on discussions with Air Force, contractor, and Navy personnel as to how much of the YJ101 development cost applied to the F404.

ANALYTICAL TECHNIQUES

This study used ordinary least squares regression as the basic tool for developing estimating relationships.[3] Before applying statistical methods, we established criteria for selecting the explanatory variables and for CER selection. The subsections that follow discuss these criteria and the statistical methods that were applied to the data collected.

Selection of Explanatory Variables

Two criteria were established before a variable was tested for significance:

1. The variable had to be logically related to cost.
2. The variable had to be known with a fair degree of accuracy during the concept formulation phase.

The search for suitable explanatory variables began with the hypothesis that the cost of an engine is a function of (1) the size of the engine, (2) the level of technology/performance incorporated into the engine, and (3) the time during which the engine is developed and produced.

Inasmuch as engine size affects the amount of raw materials that go into the engine, the size of the test facilities, and the size of the machines used in manufacturing, it appears reasonable to expect large engines to cost more than small ones. Technical complexity or difficulty is also a cost driver. The predominant cost of an aircraft

---

[3] See N. Draper and H. Smith, Applied Regression Analysis, 2nd ed., John Wiley and Sons, New York, 1981.

engine development program is incurred in achieving acceptable engine reliability and durability levels. The primary method for achieving the desired levels is full scale testing, which becomes increasingly more expensive and complex as the technology becomes more sophisticated. In the production phase, high technology and performance level translate into exotic materials and sophisticated manufacturing techniques, both of which drive up production costs.

A variable that captures the time trend was thought to be important because engine complexity and the required number of test hours--hence, cost--have increased and are continuing to increase over time. Thus, we hypothesized, an ideal CER should have at least three explanatory variables: one that is indicative of engine size, one measuring the level of technology/performance, and one reflecting the time frame during which the engine is developed or produced.

This hypothesis and the two criteria described above limit our choice of independent variables to those shown in Table 4. Statistical techniques described in the next section were applied to these variables.

Of course, engine characteristics alone cannot explain variability in program costs. Schedule, production rates, commonality between engine types, management, funding, state-of-the-art advance, availability of labor, and investment in capital tools all affect costs but cannot be captured in a simple model. A parametric cost model based on data from a wide assortment of programs is not sensitive to small changes, and it assumes that every program will have its fair share of technical, programming, and funding problems. Only when an explanatory

Table 4

EXPLANATORY VARIABLES TO BE TESTED FOR STATISTICAL SIGNIFICANCE

Size	Performance/Technology	Time
Thrust[a]	Turbine inlet temperature[a]	Time of Arrival[a]
Weight	Thrust to weight ratio[a]	
Airflow	Mach number[a]	
	Total pressure	
	Specific fuel consumption	
	Thrust per pound of airflow	

[a]These variables may be easier to obtain in a long range planning study than the others.

variable demonstrates a consistent and perceptible influence on a variety of programs can it be included in a cost model.

Limiting the Number of Explanatory Variables

It is a generally accepted maxim among analysts and statisticians that the fewer independent variables a model has, the better it will stand the test of time. There are practical and theoretical reasons for this view. The most important are:

1. Models with too many variables usually result in large prediction variances because many parameters have to be estimated.

2. Multicollinearity is more likely to occur with a large number of variables.

When comparing competing models, therefore, we favored those that had the fewest explanatory variables while maintaining predictive quality.

Criteria for CER Selection

The development of potentially useful CERs requires the selection of the "best" equation. Our interpretation of "best" is predictive capability rather than statistical quality. Realizing predictive capability requires more than fitting a line to a collection of data points; the selected independent variables should be key measures of underlying trends and not be overly influenced by one or a few data points. For example, in several cases some independent variables slightly improved the statistical properties of a few equations. However, they were eliminated from consideration because they failed to meet the above criteria.

CER selection criteria also required the signs of the coefficients in the equations to be consistent with intuitive notions of what constitutes more technologically advanced achievement with time: positive coefficients on variables that are more difficult and hence more costly to achieve, and negative coefficients on variables for which smaller values are more difficult to achieve. For example, one would expect rising turbine temperature to increase costs, and this variable does have a positive coefficient in the derived equation. In all cases the equations satisfied these criteria.

Statistical Analysis

In this Note we consider a CER as a regression equation that represents a relationship of the form

$$Y = X \beta + \varepsilon ,$$

where Y is a vector of size n, X is an n by p matrix of explanatory variables, β is a vector of p regression coefficients (including an intercept), and ε is an error vector of size n. Predicted or estimated values are indicated by a superscript *, as in β^* to denote an estimate of the β vector of true regression coefficients. Lower case letters are used for individual matrix elements, such as element y_i of vector Y.

After identifying those engine characteristics expected to drive engine costs, we computed all possible regressions and screened them for various numbers of select independent variables. For each dependent variable, the regression models with the smallest estimating errors were considered candidates for further evaluation. Just accepting the one model with the smallest error was not appropriate, because some models with small errors were invalid for other reasons. The number of models actually evaluated was not the same for all cases. We always selected a set of models large enough to be reasonably likely to include the overall best model (along with at least a few other potentially useful models). The models dropped from the analysis at this point were of necessity inferior to those that remained. The Mallows' C(p) statistic,[4] a measure of the total squared error (bias plus random),

[4] See C. Mallows, "Some Comments on C(p)," Techometrics, 15 (1973), pp. 661-675.

and the multiple correlation coefficient were used to help cull out those combinations of characteristics that showed potential for use as independent variables.

The next step was to compute full sets of statistics for the candidate models. The F statistic for the model as a whole and a t-statistic for each estimated coefficient were computed to test for statistical significance at the 10 percent level. Models that did not test out as statistically significant or displayed nonrandom residual patterns were dropped from the analysis.

Three statistical criteria were used to evaluate the remaining CERs: the total mean square error of the estimate, the influence of individual data points or sets of points, and collinearity. Estimation error is usually measured by the variance of the estimate, because we normally deal with unbiased estimators. It is possible, however, for a biased estimator to have a lower total mean square error than an unbiased estimator if its variance is low enough. Additionally, sometimes one or a few individual data observations drive the values of the estimated parameters in a CER more strongly than does the rest of the sample. Indeed, even the set of variables that are statistically significant may be largely determined by only one or a few observations. Such results are obtained when the data sample is not homogeneous, and the CER then represents effects that are not typical of the sample as a whole, but rather are associated with a peculiar subset of observations. In some cases this subset can provide useful information. In others, the peculiar observations are bad data that should be dropped from the analysis so that a useful CER can be developed from the remaining data.

Finally, collinearity can increase the variances of estimated coefficients, producing large prediction intervals or masking the validity of a test of statistical significance.

A number of diagnostic statistics have been developed to evaluate these potential problems. A thorough discussion of several of them is presented in a recent book by Belsley, Kuh, and Welsch.[5] Most of the diagnostics used in this evaluation are taken from that source. Some additional helpful material was found in a paper by R. R. Hocking.[6]

As mentioned above, a useful estimate of a standardized measure of total mean square error is the C(p) statistic:

$$C(p) = \frac{\text{Residual Sum of Squares}}{(\text{Standard Error of Estimate})^2} + 2p - n$$

When a large number of alternative models are being considered, C(p) provides an easily computed criterion for comparing equations. Several statistics that can help identify influential observations are listed in Table 5. Formulas for computing most of them are given by Belsley.[7] $DFBETAS_{ij}$ measures the influence of the ith observation on the jth estimated regression coefficient. A large value, greater than $2/\sqrt{n}$, indicates that the ith observation has strong influence on the jth coefficient.

A similar measure is $DFFITS_i$, which measures the influence of the ith observation on the fit, the estimated value of the dependent

[5] David A. Belsley, Edwin Kuh, and Roy E. Welsch, Regression Diagnostics: Identifying Influential Data and Sources of Collinearity, Wiley, New York, 1980.

[6] R. R. Hocking, "The Analysis and Selection of Variables in Linear Regression," Biometrics, 32, March 1976, pp. 1-49.

[7] Belsley (1980).

Table 5

INDICATORS OF INFLUENTIAL OBSERVATIONS

Indicator	Item Measured	Suggested Cutoff Value
DFBETAS	Change in an estimated regression coefficient caused by deletion of an observation	$2/\sqrt{n}$
DFFITS	Change in fitted value caused by deletion of an observation	$2\sqrt{p/n}$
Cook's Distance, D	Change in estimated regression coefficient vector caused by deletion of an observation	F(p, n-p, 0.50)
Hat diagonal, h	Influence of a value of dependent variable on corresponding fitted value	2p/n
Studentized Residual, RSTUDENT	Estimated normalized residual	2.0
COVRATIO	Sensitivity of covariance matrix to deletion of an observation	$1 \pm (3p/n)$

variable. Values larger than $2\sqrt{p/n}$ denote influential cases. A related measure is Cook's distance, D, which measures the change in the entire β vector because of deletion of an observation.[8] It can be

[8] See R. Cook, "Detection of Influential Observations in Linear Regression," Techometrics, 19 (1977), pp. 15-18.

evaluated by comparing it with an appropriate F distribution.

Another useful statistic is h_i, the diagonal element of the hat matrix:

$$H = X(X^tX)^{-1}X^t$$

where X^t is the transpose of X, and -1 indicates matrix inverse. Each h_i reflects the influence of an observed data point y_i on the fitted value y_i^*. A recommended cutoff value is 2p/n; values higher than this indicate leverage points that may have undue influence on the fit.

The studentized residual, RSTUDENT, for observation i is standardized using the hat matrix element h_i and the estimated error variance for the case with i excluded s(i):

$$\text{RSTUDENT}_i = \frac{e_i}{s(i)\sqrt{1-h_i}}$$

where e_i the straightforward residual $y_i - y_i^*$. Typically, one might select a magnitude of 2.0 for RSTUDENT as a screening criteria. Observations with larger values would then be considered outliers.

Another way to note the effect of a single observation is to compare covariance matrices of the estimated coefficients with the observation and without it. The parameter COVRATIO is defined as the ratio of the determinant of the covariance matrix with the observation deleted to the determinant of the full covariance matrix. This ratio can be shown to be

$$\text{COVRATIO} = \left[\left(\frac{n-p-1}{n-p} + \frac{\text{RSTUDENT}_i^2}{n-p}\right)^p (1-h_i)\right]^{-1}$$

COVRATIO is useful because it measures changes in the regression coefficient variances, which can be large even when neither high leverage nor large residuals exist alone. The matrix can be considered insensitive to observations for which COVRATIO takes on values within 3p/n of 1.0. A value greater than 1 + (3p/n) indicates that dropping the observation increases the mean square error; values less than 1 - (3p/n) identify observations that would decrease mean square error if they were dropped.

These are, of course, not all of the measures that can provide useful information about influential observations. They do, however, address a variety of concerns and have been effective in revealing new information about the CERs.

For a regression coefficient vector β given by

$$\beta = (X^tX)^{-1} X^tY$$

where X and Y are the independent variable data matrix and dependent variable vector, one has associated with each coefficient β_i an eigenvalue v_i of the (X^tX)matrix. A condition index c_i can then be computed for each coefficient as

$$c_i = \sqrt{v_{max}/v_i}\ , \ (1 \le i \le p)$$

where v_{max} is the largest v_i. Any condition index larger than about 30 indicates collinearity that deserves further attention.[9] Because at

[9] Belsley (1980).

least two independent variables must be involved when collinearity exists, degradation can be a problem only when a large condition index is associated with a large proportion of the variance of two or more coefficients. Variance proportions of 0.5 or more are considered large for this evaluation. Thus, the condition index and a decomposition of regression coefficient variances provide an indication of potential collinearity problems.

III. ESTIMATING RELATIONSHIPS

This section gives the results of applying our analyical procedure to the engine technical and cost data bases discussed in the previous section. Equations are presented for the (1) cost of development of an engine to the MQT; (2) total development cost, which includes MQT and all subsequent product improvement during the life of the engine; (3) cumulative average unit production cost at 1000 units; and (4) Time of Arrival, the predicted date for an engine having a specific set of characteristics passing the MQT. Each equation is discussed in detail, including definitions, background information, data base, modifications to the data and results.

DEVELOPMENT COSTS

Two distinct development costs are analyzed in this study. The development cost to MQT, MQTDEVCOST, is associated with the endurance test, after which the engine is considered to be sufficiently developed for installation in a production military aircraft and is suitable for operational use in the field. MQTDEVCOST includes initial design, engineering, prototype tooling, materials and fabrication, and assembly and testing of components and complete engines. Not included are any costs associated with demonstrator programs that may have provided the basic technology required to develop the new engine or any costs of production tooling associated with the procurement phase. No flight test engines are included.

The cumulative cost of development of all series of a particular engine model through some number of production engines is designated

TOTDEVCOST, which includes the expenses involved in developing a new engine to MQT, as outlined above, plus the costs to correct service-related deficiencies, and costs for continued performance and reliability improvement over time. A performance-improved model must pass an additional endurance test at the higher performance level. The cost of continued development beyond MQT for engines that are in production over several years can exceed the cost of development up to MQT. This is illustrated in Fig. 1, which plots percentage of cost to MQT versus percentage of time to MQT for eight engine programs.

Development Cost to MQT

In addition to deriving a CER based on the criteria presented above, we present a discussion of the diagnostic statistics for the derived model. This example should aid in interpretation of the diagnostic statistics for the other models, which are displayed in the appendix.

Cost Estimating Relationship. The preferred equation and relevant statistics for development costs to MQT are shown in Table 6. The 16 data points used to derive this equation represent 16 separate development programs.[1]

The estimating relationship has three explanatory variables: maximum thrust, Mach number, and turbine inlet temperature. These variables have intuitive appeal as well as statistical significance.

Maximum thrust can be considered a measure of the physical size of the engine. More than half of the cost of developing an engine is for

[1] Before analysis, we dropped one engine, the TF41, from the engine data base. This engine is an advanced version of a British engine, the RB.168.25. Its development is better characterized as a performance upgrade rather than a full development.

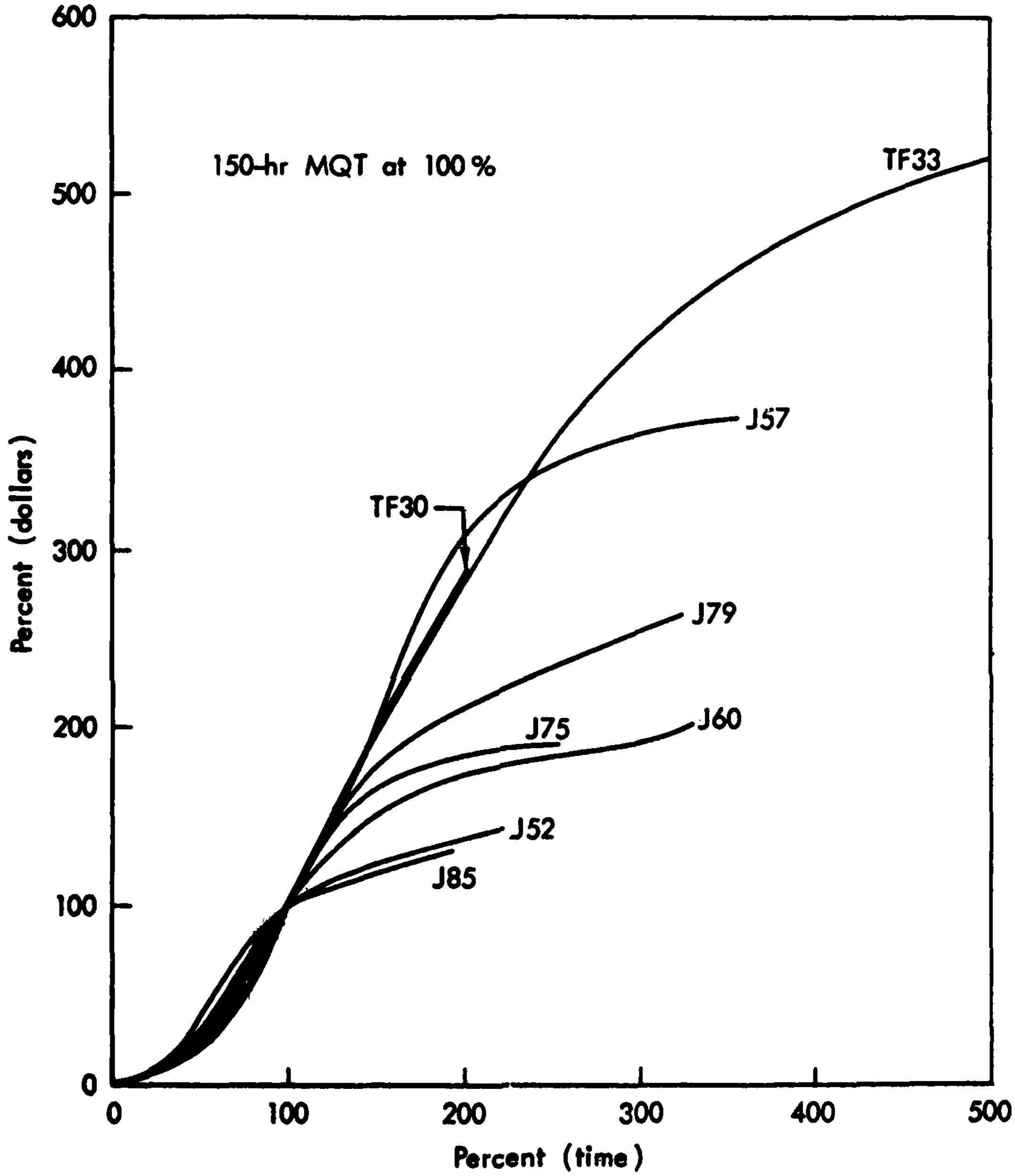

Fig. 1--Total development cost versus time for selected aircraft turbine engines

SOURCE: J. R. Nelson and F. S. Timson, Relating Technology to Acquisition Cost: Aircraft Turbine Engines, The Rand Corporation, R-1288-PR, March 1974.

Table 6

AIRCRAFT TURBINE ENGINE DEVELOPMENT COST TO MQT
(16 turbojet and turbofan engines)

$$MQTDEVCOST^{a} = \underset{}{-845.804} + \underset{(.074)^{b}}{.005\ (THRMAX)} + \underset{(.000)}{249.838\ (MACH)} + \underset{(.001)}{0.313\ (TEMP)}$$

R^2 = .93
SE = 84.7
F = 54.6

Note: MACH has a value of 1 for engines designed for subsonic flight.

[a]Millions of FY 1980 dollars.

[b]Level of significance.

test hardware; and as an index of engine size, thrust reflects the cost of hardware. Engine development programs will use as many as 50 test engines and equivalent spares.

The Mach number can be considered an indicator of the environment in which the engine must operate, and the operational environment is a strong determinant of the amount of testing required. More than one-quarter of the cost of the development program is associated with testing.

Turbine inlet temperature is the most important variable in the analysis. One obvious explanation for the statistical significance of turbine inlet temperature, even after the major performance parameters have been included in the equation, is that temperature plays a dominant

role in the thermodynamics of engines; and consequently, a major development goal has been ever-higher temperatures. These higher temperatures have been one of the chief sources of improved performance as measured by the major performance parameters. Although this by itself is important, the fact that turbine temperature serves as a proxy for many major and minor parameters, as well as for material content, should not be overlooked. Because turbine inlet temperature is also highly correlated with time ($R^2 = .9$), it serves as a substitute for a time term. (An earlier Rand study shows the turbine inlet temperature has increased at an average rate of about 35 to 40 deg. R per year.)[2]

The results for this equation are displayed graphically in Fig 2. The 45-degree line represents the average trend or expected value for MQTDEVCOST over the period. The points represent the calculated (predicted) versus actual costs for the engines in the data base. The calculated MQTDEVCOST, which is determined by inserting an engine's characteristics into the equation, is plotted on the vertical axis. The horizontal axis shows the actual cost through MQT. The scatter of the residuals about the 45-degree line does not appear to violate any assumption usually made about the distribution of errors.

Regression Diagnostics. Regression diagnostics have helped us in many ways. First, they have flagged improperly transcribed data, as well as inaccurate or incomplete data. Second, they have confirmed our belief that several engines--F100, TF39, J58--are sufficiently different from the others that their inclusion adds valuable information and broadens the applicability of the model.

---

[2] J. R. Nelson et al., "Future V/STOL Airplanes: Guidelines and Techniques for Acquisition Program Analysis and Evaluation," The Rand Corporation, N-1242-PA&E, October 1979.

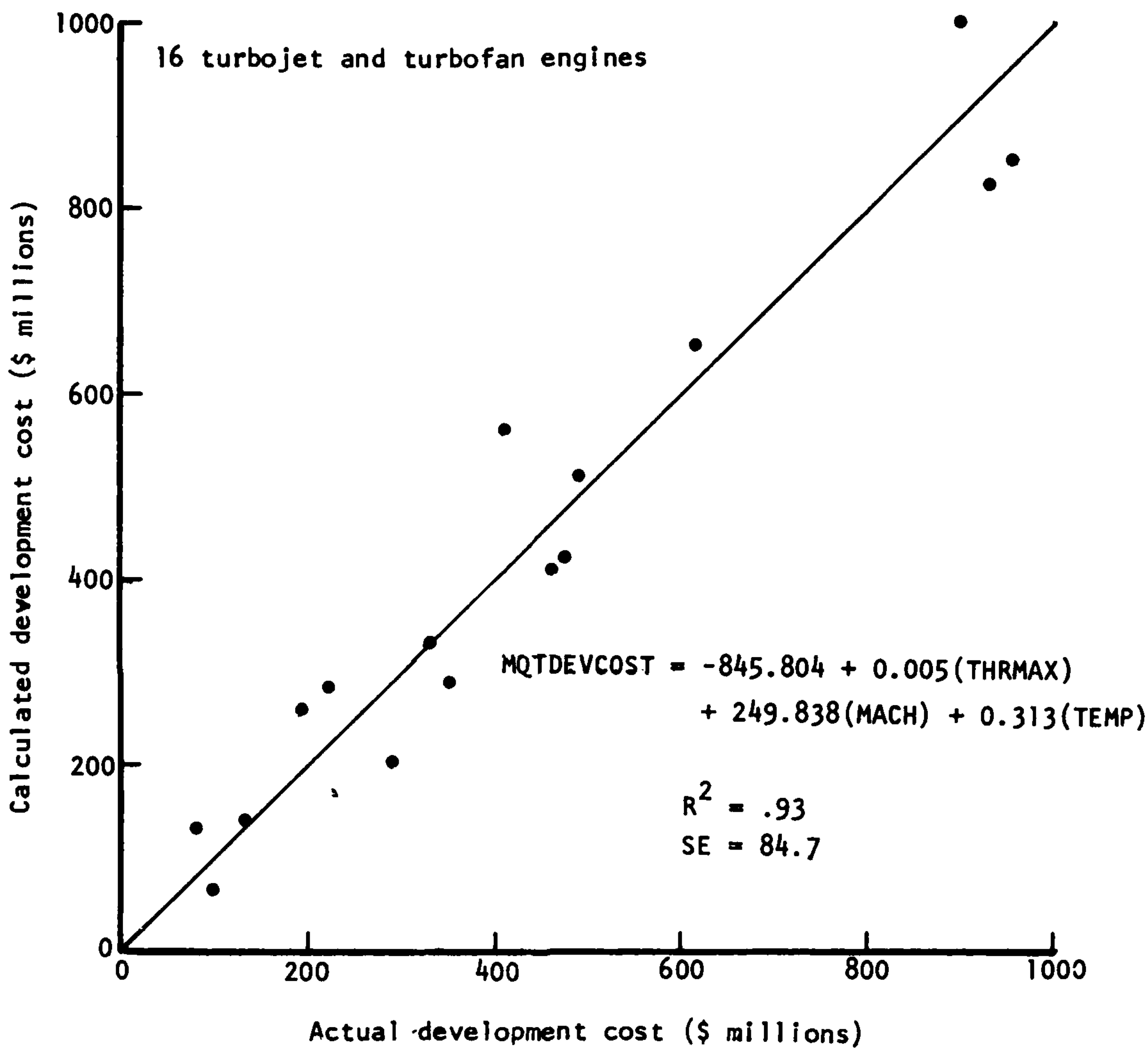

Fig. 2--Development cost (MQT), preferred equation

In the discussions that follow, regression diagnostics are used to identify data points that have a disproportionate influence on the MQTDEVCOST model and to determine which elements of the model are influenced most by these data points. In determining cutoff values, we use rules of thumb suggested by Belsley, Kuh, and Welsch.

We begin the discussion of regression diagnostics with the condition index. The maximum condition index for this equation is 21.7, which is below the established cutoff value of 30. Although a small condition index indicating degrading collinearity is absent, the presence of a high condition index alone is a necessary but not sufficient condition to reject a model. Most analysts also require high variance proportions (say, greater than 0.5) for two or more estimated regression coefficient variances. When both the condition index and variance proportions exceed these cutoff values, they provide a measure of the degree to which the model has been degraded by collinearity.

Table 7 shows studentized residuals, hat diagonal, covariance ratio, DFFITS, DFBETAS, and Cook's distance values for each data point used in the regression. (Statistics that approach or exceed their cutoff values are underlined.) Before beginning the analysis we expected the F100, as well as the TF39 and J58 engines, to be flagged by the diagnostic statistics. The F100 should be an influential observation because it is the most technically advanced engine in the sample, having the highest turbine inlet temperature and thrust to weight ratio. The TF39 should also be identified as an influential observation because, although it is a subsonic engine, it has the highest thrust rating of all the engines in the sample. The large thrust output is due to its

Table 7

REGRESSION DIAGNOSTICS FOR MQTDEVCOST EQUATION

Engine	Residual	Hat Diagonal	Covariance Ratio	DFFITS	DFBETAS Intercept	DFBETAS THRMAX	DFBETAS MACH	DFBETAS TEMP	Cook's D
F100	1.6098	0.2799	0.8439	1.0037	-0.7694	-0.3440	0.3494	0.6645	0.222
F101	1.5470	0.2441	0.8526	0.8791	-0.5387	0.0824	0.0038	0.4850	0.173
F404	-0.5172	0.2857	1.8011	-0.3271	0.2544	0.2036	-0.0135	-0.2795	0.028
TF30	-2.1377	0.0998	0.3920	-0.7120	-0.0225	0.0019	-0.4270	0.1101	0.098
TF33	-0.1444	0.2300	1.8255	-0.0789	-0.0593	-0.0415	0.0454	0.0392	0.002
TF34	-0.9538	0.3640	1.6206	-0.7216	0.3031	0.3922	0.4100	-0.5413	0.131
TF39	-0.4426	0.7001	4.4006	-0.6762	-0.0924	-0.4758	0.4509	0.0095	0.123
J52	0.7383	0.1335	1.3470	0.2898	0.1369	-0.0687	0.1069	-0.1147	0.022
J57	1.1477	0.1146	1.0175	0.4128	0.2705	-0.0072	-0.0669	-0.1670	0.042
J58	-1.6939	0.4651	1.0477	-1.5796	0.1563	-0.5458	-1.1382	0.3574	0.540
J60	0.3775	0.2061	1.6945	0.1923	0.0561	-0.0843	-0.0828	0.0167	0.010
J65	-0.5646	0.1430	1.4741	-0.2306	-0.1300	0.0383	0.0755	0.0548	0.014
J71	-0.8172	0.0956	1.2371	-0.2656	-0.1093	0.0713	0.0221	0.0414	0.018
J75	0.7089	0.3004	1.6932	0.4645	0.3432	0.3226	0.1176	-0.3906	0.056
J79	0.6059	0.1225	1.4155	0.2264	0.1196	0.0333	0.1094	-0.1311	0.014
J85	-0.0043	0.2155	1.8054	-0.0023	-0.0003	0.0013	-0.0012	0.0002	0.000
CUTOFF VALUE	±2.0	.5	(.3,1.8)	±1.0	±.5	±.5	±.5	±.5	.9

Note: Statistics that approach or exceed their cutoff values are underlined.

large size and the fact that much of its thrust is generated in a mode quite different from the other engines in the sample. Also, it is the only large, transport type of engine in the data base. The third expected outlier is the J58. This engine is the only engine in the sample designed for a high altitude, high speed reconnaissance mission, which requires a considerably different design and testing approach.

A routine analysis of residuals only identified the TF30 as being an outlier. The diagnostic statistics do a better job in this regard (see Table 8). This table shows those engines identified as being influential by the regression disgnostics: An X indicates that the suggested cutoff value for a given statistic (column heading) is approached or exceeded by a particular engine (row heading). For many statistics the F100 is flagged as being an important data point in influencing the regression *coefficients*. In addition, several of the regression diagnostics identify the TF39 as being potentially different from other engines in the sample. The J58 has the highest Cook's distance value and conspicuously exceeds the DFFITS cutoff value and approaches the limit established for the hat matrix diagonal. Two of the four DFBETAS cutoff values are exceeded as well. The results of each diagnostic are discussed individually below.

Table 8

ENGINES IDENTIFIED AS BEING INFLUENTIAL BY THE REGRESSION DIAGNOSTICS
(MQTDEVCOST Equation)

Engine	Residual	Hat Diagonal	Covariance Ratio	DFFITS	DFBETAS Intercept	DFBETAS THRMAX	DFBETAS MACH	DFBETAS TEMP	Cook's D
F100	--	--	--	x	x	--	--	x	x
F101	--	--	--	--	x	--	--	x	--
F404	--	--	x	--	--	--	--	--	--
TF30	x	--	--	--	--	--	--	--	--
TF33	--	--	x	--	--	--	--	--	--
TF34	--	--	--	--	--	--	--	x	--
TF39	--	x	x	--	--	x	x	--	--
J58	--	x	--	x	--	x	x	--	x
J85	--	--	x	--	--	--	--	--	--

Studentized residual. An examination of the studentized residuals indicate the TF30, the first turbofan engine with an afterburner, exceeds cutoff value. The distribution of these residuals do not differ greatly from the Gaussian (normal).

Hat-matrix diagonal. Using a cutoff value of .50, two engines--the TF39 and J85--show indications of being multivariate outliers. The importance of this fact depends on the values of the other diagnostics for those engines. However, we recognize that they have unique attributes and have decided that their inclusion is necessary to make the model broadly applicable.

Covariance ratio. Using the formula given in Table 5, we calculate cutoff values outside the range of 0.2 to 1.8. Engines outside this range can be considered extreme, affecting the efficiency of coefficient estimation. Four engines are outside this range--F404, TF33, TF39, and J85. Of these, the TF39, which has already been identified as a possible influential observation, has the largest deviation from the cutoff value. The remaining three engines have values essentially equal to the higher suggested cutoff.

DFFITS. DFFITS shows how the regression coefficient will change when the case is deleted from the model. This diagnostic identifies two engines, F100 and J58, both of which have been hypothesized as potentially influential observations.

DFBETAS. This diagnostic measures the influence of each data point on the individual coeffficients. It is clear from the DFBETAS that the F100, TF39, and J58, along with the F101, are very influential data points, as expected.

Cook's distance. Cook's distance provides a method for examining the change in the estimate of β relative to the usual confidence measures when a single case is deleted. Again, the F100 and J58 have the largest values.

As we initially expected, the F100, TF39, and J58 proved consistently to be the most influential engines in this model. Because future engines are expected to be more like these than the others, these engines have remained in our sample.

Unlike many CERs, this equation is linear rather than exponential (log-linear). Our preference for a linear equation comes from higher overall statistical qualities in the traditional measures--F-test, level

of significance of the independent variables (t-test), and coefficient of determination (R^2)--as well as for the regression diagnostics and residual plots. Indeed, in the log-linear format some independent variables tested as not significant, and the condition index exceeded our criteria, indicating a degree of collinearity.

Total Development Cost

In estimating total development costs the preferred equation, in addition to thrust and Mach number, includes a production quantity term. This variable is intuitively appealing, because the amount of resources devoted to improving a particular engine model through retrofitting should be related to the quantity of engines that are produced and operating in the field as well as its technical characteristics. Temperature could serve as an additional explanatory variable but is not included here to keep the number of independent variables to a minimum.

With this equation, the independent variables used to predict, say, the costs through the 2000th unit, have values that reflect the performance/technology inherent in the 2000th engine. Thus, in the absence of any change in the performance/technology level in that engine, the total development cost reflects efforts devoted to improving reliability and durability, which is related to the quantity of engines produced. However, when performance is upgraded, the equation captures both the costs of reliability and performance improvement. The equation and a summary of its statistical properties is shown in Table 9. Figure 3 graphically displays the results.

A rule of thumb in the industry is that development costs double between MQT and 2000 engines and the data substantiate this rule. Of course, in our approach we are hoping to capture underlying trends in

Table 9

TOTAL ENGINE DEVELOPMENT COST
(29 turbojet and turbofan engines)

$$\underset{[a]}{\text{TOTDEVCOST}} = -525.763 + \underset{(.000)^{b}}{0.023\ (\text{THRMAX})} + \underset{(.000)}{401.022\ (\text{MACH})} + \underset{(.002)}{0.070\ (\text{QTY})}$$

R^2 = .79
SE = 196.1
F = 31.0
Durbin Watson = 1.9

[a] Millions of FY 80 dollars.

[b] Level of significance.

the development process, with the expectation such trends will continue. Development costs do continue after initial qualification of an engine. Some allowance for these, even though imprecise, is essential in financial planning.

The engines that exceed the established cutoff criteria for the diagnostic statistics are identified in Table 10. The F100, J79, and J57 are consistently influential engines. The latter two engines were produced in very large quantities and have the largest values for the QTY variable of all the engines in our sample. Far fewer F100s have been produced to date but the total development cost of that engine greatly exceeds the cost of all other engines.

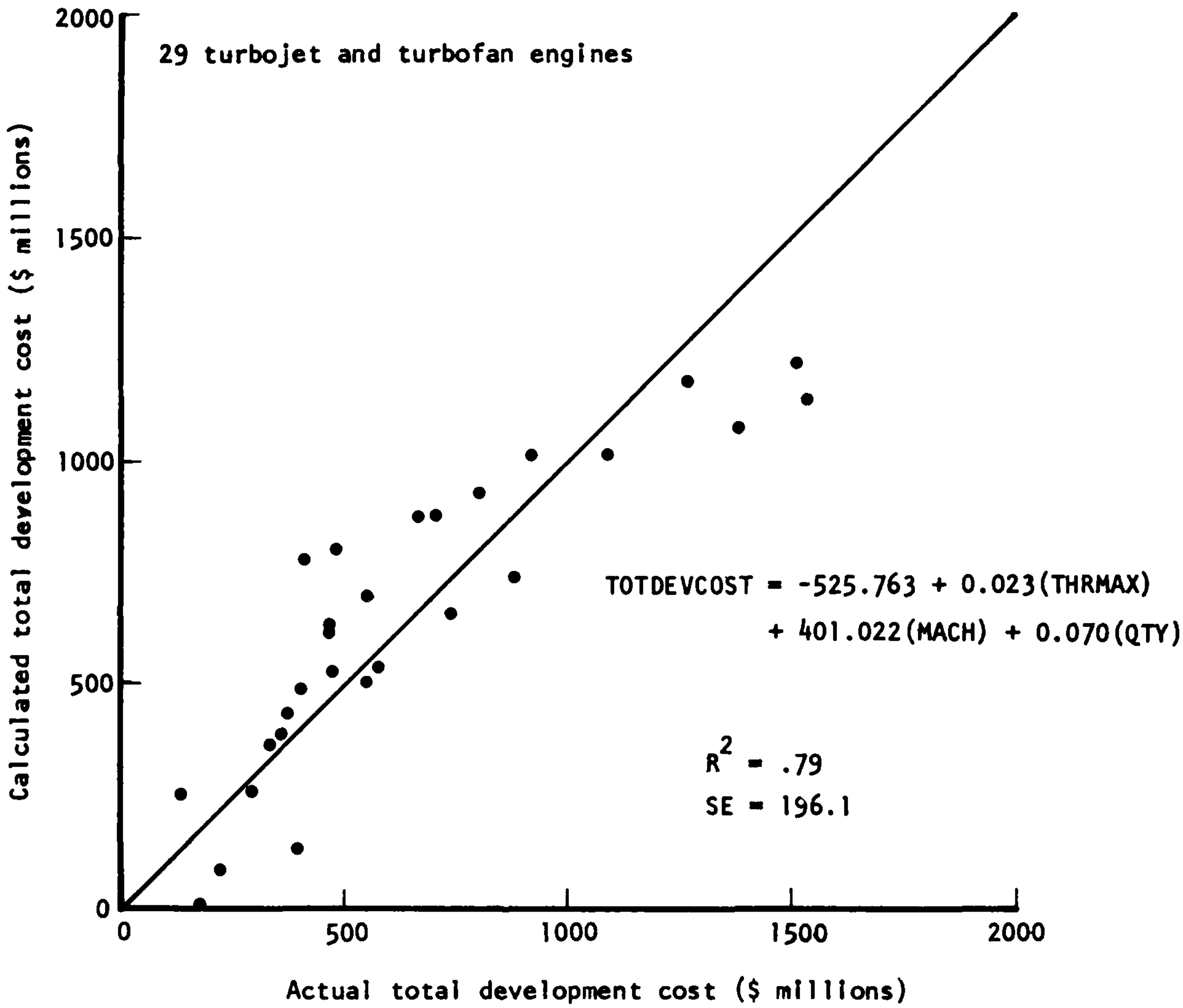

Fig. 3--Total development costs

Table 10

ENGINES IDENTIFIED AS BEING INFLUENTIAL BY THE REGRESSION DIAGNOSTICS
(TOTDEVCOST equation)

		Hat	Covariance		DFBETAS				Cook's
Engine	Residual	Diagonal	Ratio	DFFITS	Intercept	Mach	QTY	THRMAX	D
F100	x	--	x	x	x	x	--	--	--
TF30	x	--	--	--	x	--	--	--	--
TF33	--	--	x	--	--	--	--	--	--
TF34	--	--	--	--	x	--	--	--	--
J57	--	x	x	x	--	--	x	--	--
J75	--	--	--	--	--	--	--	x	--
J79	--	x	x	--	--	--	x	--	--
J85	--	--	x	--	--	--	--	x	--

A potential problem in using a series of observations for each engine to represent the cumulative costs and quantities at the end of successive years is that one of the basic assumptions of regression theory might be violated--that the errors in the successive data points are independent of each other. The failure of this condition is called serial (or auto) correlation, and its effect is to invalidate the standard error, student t, and F statistics relating to confidence measures for the equation and its coefficients. (The resulting coefficients are still maximum likelihood estimates, but the variances are understated. That is, the standard error, t, and F statistics are not as good as indicated.) In the present case, examination of the residuals and the value of the Durbin-Watson test statistic[3] indicates that serial correlation is not a problem.

The second potential difficulty inherent in this procedure is that individual engines with more observations than others will have a stronger influence on the outcome. The least squares procedure minimizes the sum of the squares of all residuals counted equally. Unfortunately, alternative procedures that were tried reduced the sample size to such a degree that the analytical outcome proved meaningless.

PRODUCTION COSTS

As stated previously, engine production costs used in this study reflect the selling price to the government adjusted to FY 1980 dollars. The production CER uses cumulative average price at the 1000th unit as a dependent variable. These are obtained from progress curves derived from historical cost and quantity data.[4] This price includes direct

---

[3] J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least Squares Regression II," Biometrika, 37, 1951, pp. 409-429, and Vol. 38, pp. 159-178.

[4] In keeping with the DAPCA formulation of engine cost, no

and indirect labor costs, material costs, tooling, technical publications, field service, G&A expense, indirect component improvement, contributing engineering and IR&D, and profit. Because of the changing nature of costs in the aircraft turbine engine industry, there is a question whether an equation based on production experience of the 1950s, 1960s, and 1970s will be able to predict costs for the 1980s. Because the intent is to use these cost equations to predict aircraft turbine engine experience in the 1980s and even 1990s, any changes that are radically different from the past--such as methods of production, plant production capacity, production rate,[5] differences in overhead rate, and changes in contract add-ons--must be reflected in any cost assessment. We believe, however, that these data represent the continuation of an evolutionary process taking place over the past 30 years and that this trend is inherently captured in the existing models.

The 1000th cumulative average unit production cost[6] is used as the dependent variables in our regression analyses and was chosen for several reasons. First, cumulative average values are easy to convert to total production cost. Second, it gives much better statistical

---

estimating relationship has been developed for production learning curve slope. DAPCA uses a default value of 0.90 for the cumulative average curve. In the data base used in this study, all but three of the 22 engines had slcpes of 0.80 or greater. A subsample of six turbofan engines had slopes ranging upward from 0.87, with an average slope of 0.93.

[5] Because production rate is not known with precision during the concept formulation, we could not include it among the explanatory variables. Nevertheless, for estimating the costs of engines in production, it has proven useful. See M. Crazur and E. McGann, "An Investigation of Changes in Aircraft Engine Production Rate," School of Systems and Logistics, Air Force Institute of Technology, WPAFB, Dayton, September 1979.

[6] This cost does not include any development monies, either pre- or post-MQT.

results than the first unit cost. Furthermore, production should be stabilized by the time several hundred units have been produced.

Engine production cost is a function of the material content of the engine. Advances in engine performance are made possible, in a large part, through lighter, stronger, and more exotic materials. Therefore, any production CER should use independent variables that reflect an engine's material content. The best equation contains thrust (size), Mach number (performance/technology), and turbine inlet temperature (technology/time). See Table 11 and Fig. 4.

Production cost is obviously a function of engine size because large engines require more material and labor than do smaller engines. The thrust variable captures this effect. Mach number is also an intuitively appealing independent variable. High Mach number engines require advanced materials because of the high temperatures and

Table 11

AIRCRAFT TURBINE ENGINE PRODUCTION COST
(22 turbojet and turbofan engines)

$$\text{PROCOST}^{a} = -\,2228.140 + \underset{(.000)^{b}}{0.043\ (\text{THRMAX})} + \underset{(.008)}{243.250\ (\text{MACH})} + \underset{(.000)}{0.969\ (\text{TEMP})}$$

R^2 = .96
SE = 183.6
F = 135.9

[a] Thousands of FY 80 dollars.

[b] Level of significance.

pressures generated throughout the engine during sustained supersonic flight. The turbine inlet temperature is perhaps the best single indicator of engine materials. Generally, a temperature of less than 1900 deg. R implies a predominantly steel engine; higher temperatures imply a high proportion of advanced materials. Because these advanced materials are more expensive to buy and more difficult and time consuming to machine, their use results in a more expensive engine.[7]

In addition to having a sound rationale, the equation meets our statistical criteria. Those engines that exceed the cutoff criteria for the diagnostic statistics are identified in Table 12. This matrix shows the J75, F404, and the TF39 as the most influential data points. The TF39 has been consistently identified as an outlier. The model considerably overpredicts production costs for the J75 engine. Higher estimated values result because its thrust rating is not consistent with other engines in the sample having a similar level of technology, as indicated by turbine inlet temperature. The J75 has one of the lowest turbine inlet temperatures and one of the highest thrust ratings. This thrust is obtained by sheer engine size, as indicated by engine weight, 5950 lb. In addition, the J75 derives from the J57 and many parts are common to both engines. When an engine has more than the usual degree of commonality, its costs may be lower than would otherwise be expected.

The F404, has a fairly high turbine inlet temperature, indicating a technically advanced engine. It is unlike the other advanced engines in the sample, however, in that it is a small engine with a high thrust

[7] T. J. Brennan et al., Cost Estimating Techniques for Advanced Technology Engines, SAE Paper 700271, April 1970.

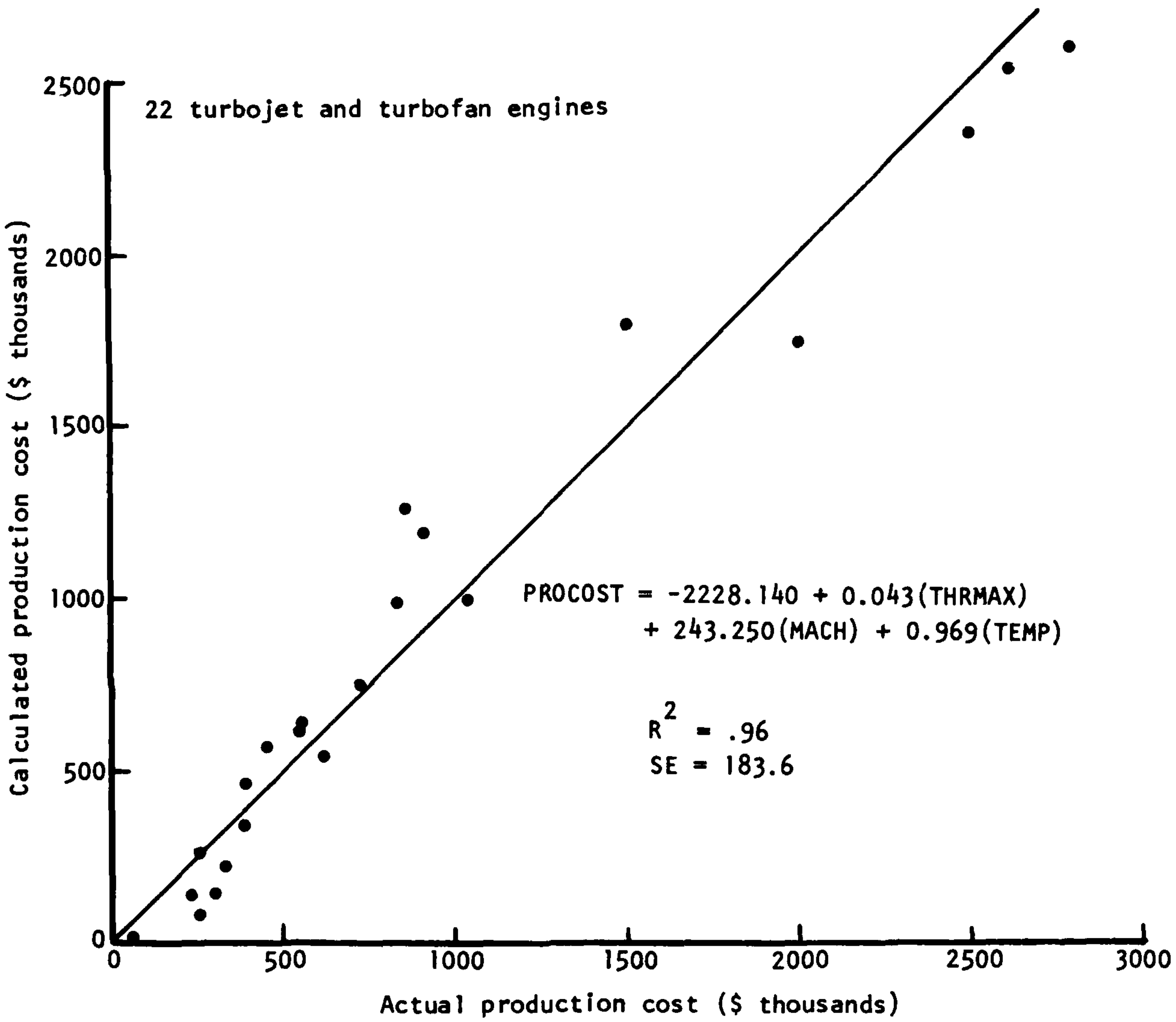

Fig. 4--Production costs, preferred equation (Cumulative average price at the 1000th unit)

Table 12

ENGINES IDENTIFIED AS BEING INFLUENTIAL BY THE REGRESSION DIAGNOSTICS
(PROCOST equation)

Engine	Residual	Hat Diagonal	Covariance Ratio	DFFITS	DFBETAS Intercept	MACH	TEMP	THRMAX	Cook's D
F100	--	--	--	--	×	--	--	--	--
F404	--	--	--	×	×	--	×	×	--
TF30	--	--	--	×	--	×	--	--	--
TF34	--	--	--	--	--	--	×	--	--
TF39	--	×	×	×	--	×	--	×	--
TF41	--	--	--	--	--	×	×	--	--
J75	×	--	×	×	×	×	×	×	×

rating. (The F404 has about the same thrust as the J79 at considerably less weight.)

All have been retained in the sample because they are valid observations that provide useful information about the nature of engine production cost and technology.

TIME OF ARRIVAL

Technology is not a directly measurable quantity, so substitute measures have been sought. One of the most successful has been the time of arrival (TOA) approach, which uses multiple regression to relate the date of an engine's successful completion of its MQT to certain of its technical characteristics. This study uses the same approach. The TOA method is provided to help the estimator discern the level of risk associated with a particular engine.

The value given by the multiple regression equation is the date when an engine with a specified set of technical characteristics is expected to pass its MQT. TOAs are measured in quarters of a year beginning with zero as the third quarter in 1942. The difference, calculated TOA - Actual TOA = Δ TOA, is the interval between the time when an engine is predicted to pass its MQT and the time it actually completes it.

The equation that best represents the military trend of technological tradeoff contains three variables (see Table 13). This equation differs from earlier equations in that all the independent variables come from the performance/technology category. Such a selection is appropriate because although TOA is a measure of time, it is a function of the performance/technology characteristics of turbine engines.

Table 13

ENGINE TIME OF ARRIVAL (TOA) EQUATION
(29 turbojet and turbofan engines)

$$TOA^{a} = -46.918 + \underset{(.000)^{b}}{8.728\ (THRWGT)} + \underset{(.001)}{0.046\ (TEMP)} - \underset{(.001)}{31.264\ (SFCMIL)}$$

$R^2 = .96$
SE = 8.6
F = 179.2

[a] In calendar quarters since the third quarter of 1942.

[b] Level of significance.

The equation contains the three most sought after characteristics in the turbine engine development process, and the sign of each coefficient is consistent with intuitive notions of what constitutes higher technological achievement. The thrust to weight ratio (THRWGT) and turbine inlet temperature (TEMP) have positive coefficients, indicating growth over time, while specific fuel consumption (SFC) is more highly valued as it is reduced. A graphical representation is plotted in Fig. 5. The 45-degree line represent the average trend or expected date of MQT over the period. Points plotted above the 45-degree line represent engines "ahead of their time"; that is, engines with characteristics yielding TOAs greater than their actual MQT dates appear earlier than predicted. Likewise, points below the line are "late" or "conservative" developments.

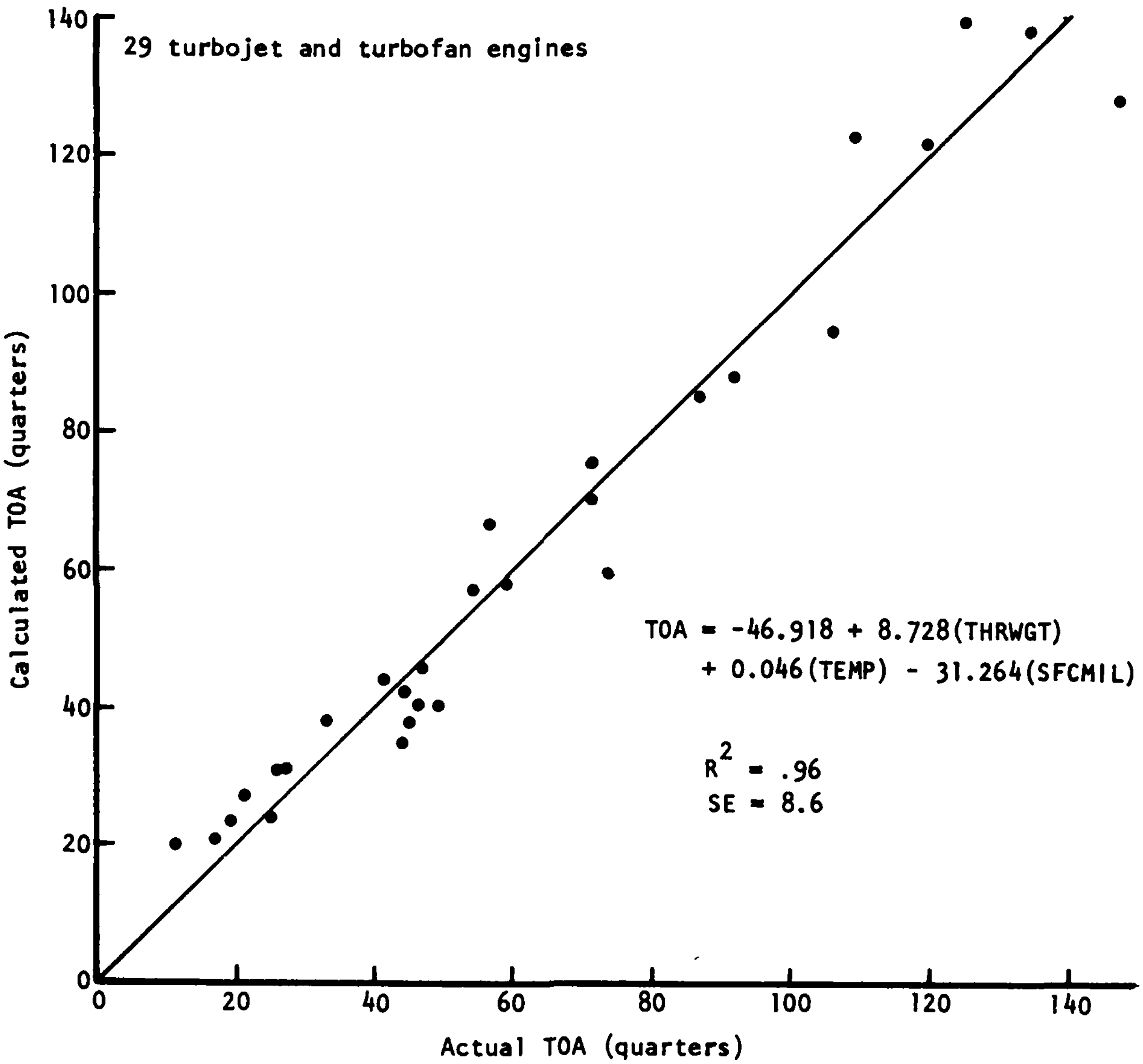

Fig. 5--TOA, preferred equation

There are many possible reasons for deviations about the trend. The equation therefore cannot be used for making fine distinctions, but if certain points or trends deviate sharply from the average, it should be possible to distinguish them. For example, two recent engines, the F100 and F404, deviate from the trend line by more than one standard error.

The model shows the F100 engine to be ahead of its time by almost 14 quarters, and that is well recognized within the propulsion community. However, the model shows the F404 to be "late" by almost 20 quarters. This tends to confirm the manufacturer and Navy's claim that they deliberately "backed off" from the ultimate performance in developing the F404, believing that a reduction of 5 to 10 percent from the theoretically possible thrust to weight ratio and specific fuel consumption would make possible a major reduction in complexity and maintenance requirements.

The F404 and the F100, along with the TF39, which was the first large turbofan engine, are the most influential data points in the TOA equation (see Table 14).

It should be stressed that these models represent the time of arrival of a demonstrated level of performance, which is assumed to represent the best efforts of the aircraft turbine engine industry; thus it is considered to be the technological state of the art of U.S. aircraft turbine engines. When an engine is predicted to be greatly advanced there is a good chance that its schedule will slip and costs will increase. Whenever the difference between the estimated and

TABLE 14

ENGINES IDENTIFIED AS BEING INFLUENTIAL BY THE REGRESSION DIAGNOSTICS
(TOA equation)

Engine	Residual	Hat Diagonal	Covariance Ratio	DFFITS	DFBETAS Intercept	DFBETAS THRWGT	DFBETAS TEMP	DFBETAS SFCMIL	Cook's D
F100	--	--	--	x	x	--	x	--	--
F101	--	--	x	--	--	--	--	--	--
F404	x	--	x	x	x	x	x	x	--
TF33	--	--	x	--	--	--	--	--	--
TF34	--	--	x	--	--	--	--	--	--
TF39	x	x	--	x	--	x	x	x	--
J52	--	--	--	--	x	--	x	--	--
J58	--	x	x	--	--	--	--	--	--
J60	--	x	x	--	--	x	--	--	--
J85	--	x	x	--	--	--	--	--	--

planned dates exceeds the standard error of the TOA equation (8.6 quarters), the development cost estimate should be adjusted. Such revised estimates are best made in the context of the program by estimators who have a sense of history and are familiar with program details.

CAUTIONS

There are several cautionary points regarding the input data. First, they must reflect the maximum capability of the engine, not the aircraft in which it is to be installed. Second, the engine performance data must be consistent in terms of the thermodynamic cycle. Third, data at growth points require some sort of forecast, a fact that has not been analyzed in the present study. The problem is to determine an improved technology level after some quantity of engines has been produced. If only a few engines are produced, technology may not be uprated. If many engines are produced, technology will probably be uprated. Fourth, predictions can be made with greater confidence when the parameter values for a new engine fall within the range of the sample data. This may occur for some new developments; however, it will probably not occur for at least the technology parameters. As a guide to the estimator, Table 15 shows the ranges of the input data for all equations presented in this report.

Table 15

MEANS AND RANGES OF INPUTS TO FINAL REGRESSION ANALYSIS

Parameter	MQTDEVCOST (16 Observations)			TOTDEVCOST (29 Observations)			PROCOST (22 Observations)			TOA (29 Observations)		
	Low	Mean	High	Low	Mean	High	Low	Mean	High	Low	Mean	High
TOA (qtr)	41	85	148	42	89	150	19	71	148	11	63	148
DEVTIME (qtr)	9	22	44	9	33	66	9	21	44	9	21	44
Development Cost[a]	82	433	961	94	629	1534	--	--	--	--	--	--
Production Cost[b]	--	--	--	--	--	--	57	914	2789	--	--	--
QTY	--	--	--	--	--	7000	--	--	--	--	--	--
THRMAX (lb)	3000	16931	40800	3000	15246	25100	1025	12845	40800	1025	11816	40800
TEMP (^{o}R)	2030	2403	3060	2060	2310	3025	1960	2296	3060	1825	2228	3060
MACH	1.0	1.8	3.2	1.0	1.8	2.5	1.0	1.5	2.5	1.0	1.5	3.2
WGT (lb)	4600	3510	7300	460	3174	5950	364	2985	7300	364	2832	7300
AIR (lb/sec)	42	294	1555	42	211	498	21	234	1555	21	238	1555
TOTPRS (lb/ft^2)	8500	23376	55650	10360	23815	52850	3400	18936	55650	1575	15679	55650
SFCMIL (lb/hr/lb thrust	0.32	0.77	1.25	0.36	0.77	1.04	0.32	0.83	1.22	0.32	0.91	1.25
THRWGT	2.3	5.1	7.9	2.4	5.3	7.9	1.7	3.6	6.5	1.7	4.0	7.9

[a] Millions of 1980 dollars.
[b] Thousands of 1980 dollars; cumulative average value at the 1000th unit.

IV. OBSERVATIONS

The statistics for the estimating relationships described in the preceding section give evidence that these relationships provide improvements in engine cost estimating capability over the DAPCA equations. Major strong points of these relationships are intuitive appeal, ease of use, fewer independent variables, and low estimating error. For the three most recent engines Table 16 shows percent deviations from observed values. Also, we have insight into the influence of one or a few engines in the data base on the derived equations. The passing of time has seen additional engine development and production, which have yielded useful data that have been added to the data base so that it represents a wider range of engine characteristics.

Table 16

PERCENT DEVIATIONS OF OBSERVED VERSUS CALCULATED VALUES

Engine	MQTDEVCOST	TOTDEVCOST	PROCOST	TOA
F100	-11	-26[a]	-6	10
F101	-12	--	-7	2
F404	6	--	+19	-14

Note: Negative values are underestimates positive values are overstimates.

[a] Through 1700 units.

The equation form selected represents a departure from the more traditional estimating relationships. Our equations are linear in both the independent and dependent variables (linear-linear form), whereas some earlier studies used logarithmic transformations. We investigated four different equation forms: linear-linear, log-linear, linear-log and log-log; each equation form has its own implication for technology and cost trends.[1] Forms other than linear-linear each had at least one of the following drawbacks:

- candidate variables were not significant
- coefficients had counterintuitive signs
- CER had large estimating error
- independent variables exhibited collinearity
- older engines exerted considerable influence on predicted values.

Thus our initial criteria and analytical results caused us to select the linear-linear format as being best suited for predicting future engine costs.

Stability of estimating relationships is revealed by the regression diagnostics. For all the preferred equations, highly influential data points were judged to have characteristics representative of future engines. They were therefore included when the final estimating relationships were developed because they are valid observations providing useful information about the nature of engine costs and technology.

---

[1] W. L. Stanley and M. Miller, Measuring Technological Change in Jet Fighter Aircraft, The Rand Corporation, R-2249-AF, September 1979, pp. 16-18.

The diagnostic statistics provide information useful in considering the effects of these influential observations upon estimated costs. Every observation with strong influence on the fit also has strong influence on the coefficients of one or more explanatory variables. The contributions of these variables to the cost or TOA estimate are especially sensitive to the influential observation. The cost estimator can tell whether a variable will be a problem in his planned application of the estimating relationship, depending upon whether the engine whose cost is to be estimated most closely resembles the influential engine or the other observations.

NOTEWORTHY ENGINES AND VARIABLES

Table 17 is a summary of the engines that deserve special note. The F100 strongly influences the coefficients in three models. The TF39 and F404 are also influential data points; however, different models are affected by individual engines to different degrees, and other engines

Table 17

NOTEWORTHY ENGINES

Dependent Variable	Statistically Influential Engines
MQTDEVCOST	F100, J58
TOTDEVCOST	F100, J79, J57
PROCOST	F404, TF39. J75
TOA	F100, F404, TF39

stand out as atypical for the different models. It should be recalled that not all engines are in the samples for all models, and this may account for some of the variation from one model to another.

Table 18 lists the explanatory variables that showed up most often in models with good statistical properties. THRMAX, MACH and TEMP were important for all the models. TOA itself is not in any of the preferred relationships. The time of arrival is thus less important now as a driver of engine costs than it was in the past, probably because of the expanded data base used for this study.

COMPARISON WITH DAPCA

Making a fair comparison with the earlier DAPCA model is not an easy task. We have benefit of more information and improved statistical methods, both of which were not available when the original DAPCA equations were derived. Nevertheless, a comparison is needed to provide a measure of progress.

Table 18

MOST FREQUENTLY OBSERVED EXPLANATORY VARIABLES

Dependent Variable	Explanatory Variable
MQTDEVCOST	MACH, TEMP, THRMAX, THRWGT, WGT
TOTDEVCOST	MACH, QTY, TEMP, THRMAX THRWGT
PROCOST	MACH, TEMP, THRMAX, THRWGT
TOA	SFCMIL, TEMP, THRWGT, THRMAX, TOTPRS

Comparing R^2 and F values of the derived equations and the DAPCA model is not useful because of different expressions of the dependent variable and differing numbers of independent variables. Therefore, we use historical simulation to compare our model with DAPCA. Historical simulation is based on the idea that if the CERs have merit now, the same relationships, derived from the then available data base, should have worked in the past. The historical simulation technique is implemented by removing the most recent engines from our data base and then performing a regression analysis on the reduced sample. The regression coefficients obtained are then used to predict the costs and TOA for the omitted engines.

This technique was used for our equation to predict costs and TOA for the F100, F101, and F404. Because the DAPCA data base did not include these engines, predictions were also obtained using the DAPCA equations directly.

Figure 6 illustrates the results of this comparison. The results confirm that DAPCA generally underestimates costs for new engines; the exception, of course, is the F404 for which DAPCA grossly overestimated development costs. The new model also tends to underestimate costs for these engines but does not experience the wide variation in estimating error that DAPCA exhibits, particularly for the development and production cost estimates. Unfortunately, total development cost data were unavailable for the F404 and F101, so a comparison was made for the total development costs of the F100 through 1700 engines. In this case, DAPCA predicts better than the new equation, even though it would not

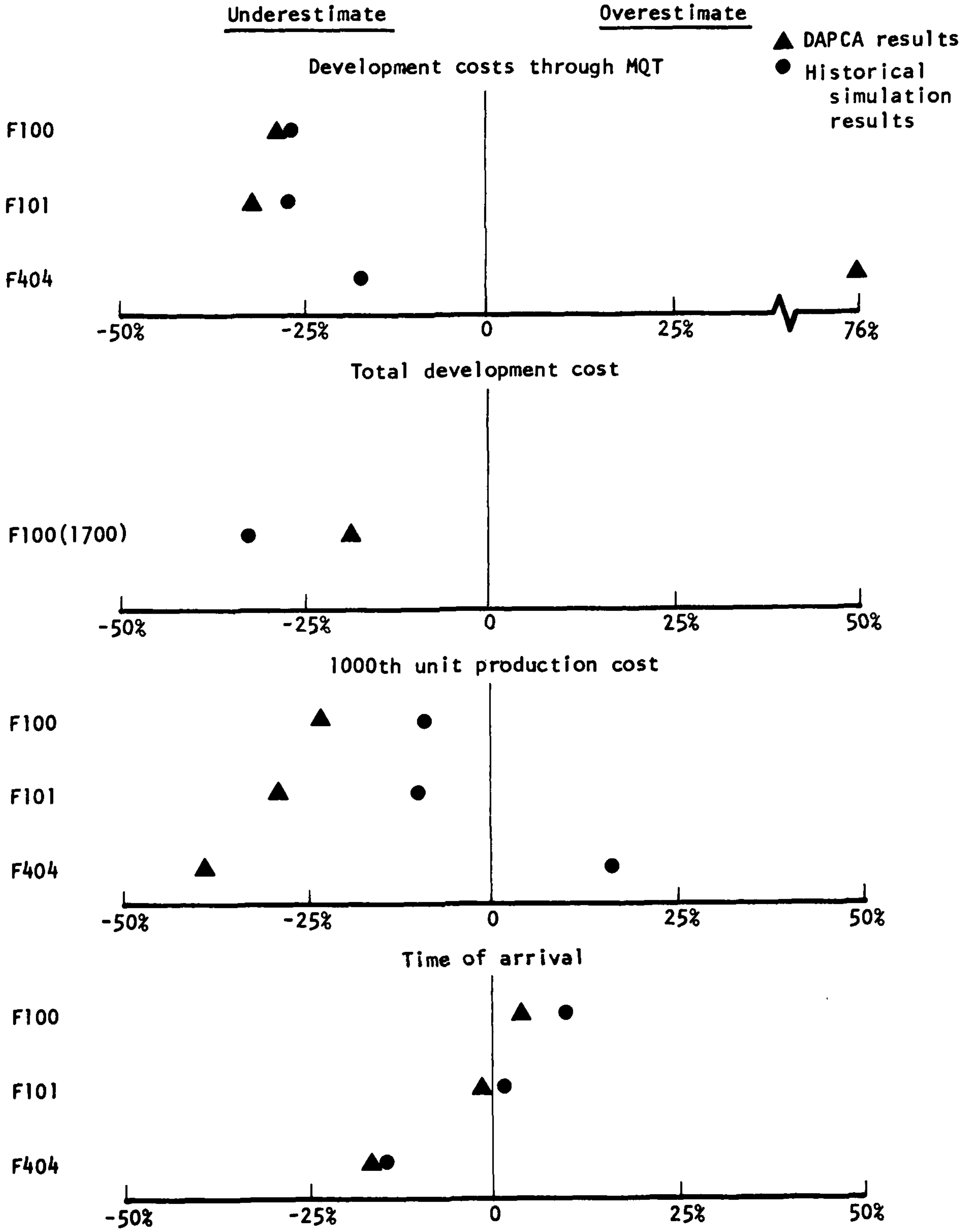

Fig. 6--Historical simulation results vs. DAPCA

have satisfied our model selection criteria. Essentially both models predict TOA equally well. In summary, then, historical simulation reveals that the new model does as well as DAPCA or better.

The new model uses few explanatory variables and meets more stringent statistical criteria than DAPCA. For the most important cost categories--costs through MQT and production--the new model predicts with less error.

IMPLICATIONS OF TOA

The TOA methodology measures the technology trend not directly, but as a time trend. It assumes that engine technology advances through steady, continuous growth. Growth in the 1980s is assumed to be equal to the growth rate experienced during the 1960s or 1970s. Various engine research and exploratory development programs of the military and the manufacturers are taken to continue at a constant rate. Clearly, if the degree of support for such work changes, or its effectiveness changes (in either direction), the estimating relationships based on past trends will lose some validity. Also, changes in acquisition procedures that may influence decisions on when to apply new technology will also affect the usefulness of these relationships. In general, meaningful use of TOA to predict risks must allow for the influence of future changes in technology development and application.

LIMITATIONS

The results described in this study are intended for estimating the cost of large, modern future aircraft engines in _the context of long-range planning studies_. Any new engine to be estimated must be

consistent with the basic assumptions under which the CERs were derived. Specifically, the CERs apply to the development and pricing practices similar to those of the 1960s and 1970s and assume basic gas turbine design will be similar to that of the engines of today. Obvious differences are evident, but no fundamental change is foreseen without stepping outside the definition of an aircraft turbine engine. Apart from the usual incremental increases in component efficiency obtained by newer alloys and novel component designs, it would appear that the estimating relationships previously discussed can be used for most of the proposed aircraft designs.

Certain designs incorporate features that, although radical in appearance, still permit the estimating relationships to be used as a base for subsequent adjustment. Swivelling nozzles, for example, are appendages that supplement the basic engine.

The lift engine is one application where difficulties will arise. Because it is considerably different in usage and design, it is doubtful whether the estimating relationship derived in this study will apply. Another case would be where the material content of the engine was radically different from those engines making up the sample--for example, an engine having cold or hot section parts largely fabricated from composite or ceramics material. Fortunately for the estimator, the engine development process has been an evolutionary one. If this trend continues, the estimating relationships derived in this study will apply to future U.S. developed and produced aircraft turbine engines.

FUTURE WORK

Data Base Expansion

Further data base expansion seems to be the key to further improvements in engine cost estimating relationships. Regression diagnostics have consistently identified the most recently developed engines as having unusually strong influence on the CERs. Therefore, the CERs should be updated as data for new engines become available. This, of course, is restricted by the limited number of engines in work at any given time, so progress will be slow. Most military turbine engines likely to be developed in the near future will be turbofans, if recent trends continue. Turbofan engines are outnumbered now in the data base by turbojets. As more turbofans are added to the data base, the total sample will become more representative of that type of engine, permitting development of estimating relationships even more useful than those obtained in this study

Parameter Interrelationships

Further investigation of potential explanatory variables may also be profitable. Those selected and our criteria greatly affected the structure of the derived models. It is unlikely that everything has been done that can be done in this area. One avenue that may be fruitful is the study of the interrelationships of the explanatory variables. Thrust, Mach, and temperature are one important set. Understanding their relationships to each other more accurately should help to point the way to new approaches to studying their influence on cost.

These areas for further improvement of engine cost estimating are really extensions of the work done in this study. Just as improvements were achieved during this present effort, so should improvements result from future work, as the availability of new data will allow.

Derivative Engines

The CERs developed assume that each new engine starts from the drawing board and follows the "normal" course to MQT and subsequent production. They do not lend themselves to estimating the incremental development costs resulting from a growth or a derivative version of an earlier, similar engine. But as budget constraints and high development costs combine to limit the development of all new military engines, strong efforts have been made to encourage and capitalize on this type of evolutionary development.[2] Given that growth and derivative engines are not only a phenomena of the past, but perhaps the wave of the future, estimating tools should be developed that address the costs associated with evolutionary growth of a common engine family.

---

[2] The Engine Model Derivative Program (EMDP), managed by the ASD Deputy for Propulsion, is an example of this philosophy. This program is designed to improve performance and durability at lower costs. It exploits existing technology and applies it to a current engine model to evolve a newer model. Under EMDP, the resulting newer model must have better characteristics than the existing engine.

Appendix

REGRESSION EQUATIONS AND STATISTICS

Table A.1

REGRESSION DIAGNOSTICS FOR MQTDEVCOST EQUATION

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB>F
MODEL	3	1173246	391082	54.562	0.0001
ERROR	12	86011.897	7167.658		
C TOTAL	15	1259258			

ROOT MSE	84.662022	R-SQUARE	0.9317
DEP MEAN	433.438	ADJ R-SQ	0.9146
C.V.	19.53269		

VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR H0: PARAMETER=0	PROB > ABS(T)
INTERCEPT	1	-845.804	154.190	-5.485	0.0001
THRMAX	1	0.005338497	0.00272854	1.957	0.0741
MACH	1	249.838	38.302369	6.523	0.0001
TEMP	1	0.312816	0.075867	4.123	0.0014

COLLINEARITY DIAGNOSTICS VARIANCE PROPORTIONS

NUMBER	EIGENVALUE	CONDITION INDEX	PORTION INTERCEP	PORTION THRMAX	PORTION MACH	PORTION TEMP
1	3.746	1.000	0.0013	0.0092	0.0060	0.0009
2	0.181933	4.537	0.0181	0.6128	0.0216	0.0026
3	0.064423	7.625	0.0523	0.0131	0.9678	0.0230
4	0.007946	21.711	0.9283	0.3649	0.0047	0.9734

Table A.2

REGRESSION DIAGNOSTICS FOR MQTDEVCOST EQUATION

ID	RSTUDENT	HAT DIAG H	COV RATIO	DFFITS	DFBETAS INTERCEP	DFBETAS THRMAX	DFBETAS MACH	DFBETAS TEMP	COOK'S D
F100	1.6098	0.2799	0.8439	1.0037	-0.7694	-0.3440	0.3494	0.6645	0.222
F101	1.5470	0.2441	0.8526	0.8791	-0.5387	0.0824	0.0038	0.4850	0.173
F404	-0.5172	0.2857	1.8011	-0.3271	0.2544	0.2036	-0.0135	-0.2795	0.028
TF30	-2.1377	0.0998	0.3920	-0.7120	-0.0225	0.0019	-0.4270	0.1101	0.098
TF33	-0.1444	0.2300	1.8255	-0.0789	-0.0593	-0.0415	0.0454	0.0392	0.002
TF34	-0.9538	0.3640	1.6206	-0.7216	0.3031	0.3922	0.4100	-0.5413	0.131
TF39	-0.4426	0.7001	4.4006	-0.6762	-0.0924	-0.4758	0.4509	0.0095	0.123
J52	0.7383	0.1335	1.3470	0.2898	0.1369	-0.0687	0.1069	-0.1147	0.022
J57	1.1477	0.1146	1.0175	0.4128	0.2705	-0.0072	-0.0669	-0.1670	0.042
J58	-1.6939	0.4651	1.0477	-1.5796	0.1563	-0.5458	-1.1382	0.3574	0.540
J60	0.3775	0.2061	1.6945	0.1923	0.0561	-0.0843	-0.0828	0.0167	0.010
J65	-0.5646	0.1430	1.4741	-0.2306	-0.1300	0.0383	0.0755	0.0548	0.014
J71	-0.8172	0.0956	1.2371	-0.2656	-0.1093	0.0713	0.0221	0.0414	0.018
J75	0.7089	0.3004	1.6932	0.4645	0.3432	0.3226	0.1176	-0.3906	0.056
J79	0.6059	0.1225	1.4155	0.2264	0.1196	0.0333	0.1094	-0.1311	0.014
J85	-0.0043	0.2155	1.8054	-0.0023	-0.0003	0.0013	-0.0012	0.0002	0.000

Table A.3

REGRESSION DIAGNOSTICS FOR TOTDEVCOST EQUATION

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB>F
MODEL	3	3574504	1191501	30.998	0.0001
ERROR	25	960957	38438.269		
C TOTAL	28	4535461			

ROOT MSE	196.057	R-SQUARE	0.7881
DEP MEAN	628.207	ADJ R-SQ	0.7627
C.V.	31.20895	DURBIN-WATSON	1.896

VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR H0: PARAMETER=0	PROB > ABS(T)
INTERCEP	1	-525.763	132.433	-3.970	0.0005
THRMAX	1	0.022730	0.00557228	4.079	0.0004
MACH	1	401.022	77.641102	5.165	0.0001
QTY	1	0.070436	0.020297	3.470	0.0019

COLLINEARITY DIAGNOSTICS VARIANCE PROPORTIONS

NUMBER	EIGENVALUE	CONDITION INDEX	PORTION INTERCEP	PORTION THRMAX	PORTION MACH	PORTION QTY
1	3.312	1.000	0.0064	0.0122	0.0055	0.0311
2	0.542866	2.470	0.0057	0.0185	0.0070	0.9553
3	0.106281	5.582	0.1844	0.9095	0.0562	0.0070
4	0.039282	9.182	0.8035	0.0599	0.9313	0.0066

Table A.4

REGRESSION DIAGNOSTICS FOR TOTDEVCOST EQUATION

ID	RSTUDENT	HAT DIAG H	COV RATIO	DFFITS	DFBETAS INTERCEP	DFBETAS THRMAX	DFBETAS MACH	DFBETAS QTY	COOK'S D
F100	-0.3120	0.1407	1.3481	-0.1262	0.0619	-0.0417	-0.0615	0.0538	0.004
F100	1.7844	0.1228	0.8155	0.6677	-0.3734	0.2285	0.3469	-0.1674	0.102
F100	2.3268	0.1153	0.5899	0.8402	-0.5282	0.2830	0.4479	0.0395	0.150
TF30	-2.0651	0.0808	0.6659	-0.6125	0.1479	-0.0530	-0.2826	0.3245	0.083
TF30	0.3447	0.1054	1.2903	0.1183	-0.0680	0.0120	0.0800	-0.0189	0.004
TF30	1.6211	0.1340	0.8972	0.6377	-0.4069	0.2539	0.2863	0.1687	0.095
TF33	-0.7392	0.1637	1.2867	-0.3271	-0.2388	-0.1555	0.2628	0.1160	0.027
TF33	-0.3927	0.1987	1.4321	-0.1955	-0.1119	-0.1314	0.1653	0.0112	0.010
TF33	0.1847	0.2162	1.4937	0.0970	0.0476	0.0607	-0.0789	0.0282	0.002
TF34	0.7294	0.1275	1.2361	0.2788	0.2600	-0.0212	-0.1780	-0.1029	0.020
TF34	1.4333	0.1143	0.9571	0.5148	0.4899	-0.0469	-0.3482	-0.0983	0.064
J52	-0.1907	0.0913	1.2878	-0.0605	-0.0198	0.0379	-0.0197	0.0268	0.001
J52	-0.3121	0.0666	1.2411	-0.0834	-0.0211	0.0569	-0.0280	-0.0132	0.002
J52	-0.7959	0.0994	1.1779	-0.2644	-0.0207	0.1108	-0.0512	-0.1881	0.018
J57	0.1507	0.0768	1.2705	0.0435	0.0330	-0.0120	-0.0123	-0.0211	0.000
J57	0.2114	0.0592	1.2422	0.0530	0.0338	0.0186	-0.0339	0.0025	0.001
J57	-1.1235	0.3427	1.4592	-0.8113	-0.0881	-0.1156	0.2387	-0.7315	0.163
J60	0.8369	0.1690	1.2628	0.3774	0.3337	-0.1887	-0.1368	-0.1123	0.036
J60	0.9359	0.1521	1.2031	0.3964	0.3538	-0.2048	-0.1561	-0.0407	0.039
J75	-1.8993	0.1035	0.7501	-0.6453	0.0672	-0.4047	0.0219	0.3158	0.094
J75	-0.9597	0.1043	1.1307	-0.3276	0.0460	-0.2312	0.0242	0.1193	0.027
J75	-0.7301	0.0907	1.1859	-0.2306	0.0481	-0.1703	0.0191	0.0108	0.014
J79	-0.8295	0.0644	1.1238	-0.2176	-0.0067	0.0341	-0.0824	0.1239	0.012
J79	0.3263	0.0527	1.2211	0.0770	-0.0010	-0.0065	0.0287	-0.0350	0.002
J79	0.6926	0.0421	1.1356	0.1452	-0.0155	0.0076	0.0493	-0.0202	0.005
J79	0.5902	0.3738	1.7749	0.4560	-0.0958	0.0029	0.0390	0.4294	0.053
J85	-0.1826	0.1900	1.4455	-0.0884	-0.0126	0.0724	-0.0480	0.0259	0.002
J85	-0.5518	0.1724	1.3527	-0.2518	-0.0179	0.2184	-0.1404	-0.0354	0.016
J85	-1.0214	0.2295	1.2888	-0.5574	0.0055	0.4240	-0.2639	-0.2963	0.078

Table A.5

REGRESSION DIAGNOSTICS FOR PROCOST EQUATION

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB>F
MODEL	3	13743738	4581246	135.855	0.0001
ERROR	18	606990	33721.690		
C TOTAL	21	14350728			

ROOT MSE	183.635	R-SQUARE	0.9577
DEP MEAN	914.000	ADJ R-SQ	0.9507
C.V.	20.09132		

VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR H0: PARAMETER=0	PROB > ABS(T)
INTERCEP	1	-2228.140	309.448	-7.200	0.0001
THRMAX	1	0.043007	0.005876019	7.319	0.0001
MACH	1	243.250	81.149876	2.998	0.0077
TEMP	1	0.968842	0.157582	6.148	0.0001

COLLINEARITY DIAGNOSTICS VARIANCE PROPORTIONS

NUMBER	EIGENVALUE	CONDITION INDEX	PORTION INTERCEP	PORTION THRMAX	PORTION MACH	PORTION TEMP
1	3.684	1.000	0.0011	0.0102	0.0064	0.0008
2	0.243189	3.892	0.0109	0.5277	0.0160	0.0017
3	0.066085	7.466	0.0429	0.0138	0.9728	0.0201
4	0.006785	23.301	0.9451	0.4483	0.0048	0.9773

Table A.6

REGRESSION DIAGNOSTICS FOR PROCOST EQUATION

ID	RSTUDENT	HAT DIAG H	COV RATIO	DFFITS	DFBETAS INTERCEP	DFBETAS THRMAX	DFBETAS MACH	DFBETAS TEMP	COOK'S D
F100	0.9893	0.3010	1.4373	0.6492	-0.4660	-0.1615	0.3104	0.3550	0.105
F101	1.2513	0.2581	1.1910	0.7381	-0.3956	0.1232	0.1304	0.3022	0.132
F404	-1.9437	0.2622	0.7634	-1.1588	0.9799	0.6561	-0.2208	-0.9527	0.291
TF30	1.7097	0.2061	0.8393	0.8712	-0.1268	0.0909	0.6920	-0.0991	0.171
TF33	-0.1348	0.1804	1.5269	-0.0632	-0.0479	-0.0451	0.0299	0.0369	0.001
TF34	-1.0646	0.2726	1.3347	-0.6516	0.3797	0.3621	0.3321	-0.5407	0.105
TF39	0.7838	0.6531	3.1432	1.0753	0.2138	0.8181	-0.6015	-0.1204	0.295
TF41	-1.6970	0.1827	0.8224	-0.8022	0.3222	0.1874	0.5347	-0.5378	0.146
J33	1.0342	0.1037	1.0986	0.3517	0.1608	-0.0662	-0.1296	-0.0549	0.031
J35	0.9715	0.1002	1.1254	0.3243	0.1135	-0.0859	-0.1312	-0.0114	0.026
J47	0.5989	0.0957	1.2786	0.1949	0.0551	-0.0539	-0.0888	0.0085	0.010
J48	-0.0092	0.0916	1.3836	-0.0029	-0.0013	0.0002	0.0014	0.0004	0.000
J52	-0.4045	0.1020	1.3471	-0.1364	-0.0491	0.0098	-0.0831	0.0580	0.005
J57	0.4058	0.0672	1.2964	0.1089	0.0695	0.0219	0.0044	-0.0558	0.003
J60	0.4231	0.1094	1.3533	0.1483	0.0232	-0.0649	-0.0591	0.0239	0.006
J65	0.2587	0.0729	1.3345	0.0725	0.0393	-0.0016	-0.0159	-0.0217	0.001
J69	0.4242	0.1235	1.3748	0.1592	0.0332	-0.0746	-0.0495	0.0143	0.007
J71	-0.4765	0.0536	1.2595	-0.1134	-0.0354	0.0128	-0.0194	0.0228	0.003
J73	-0.4512	0.1206	1.3628	-0.1671	-0.0567	0.0079	-0.1139	0.0740	0.007
J75	-3.5305	0.3395	0.2109	-2.5309	-1.7098	-1.9258	-0.9091	2.0694	0.978
J79	0.2320	0.1293	1.4253	0.0894	0.0373	0.0283	0.0581	-0.0516	0.002
J85	-0.4128	0.1748	1.4636	-0.1900	0.0078	0.0948	-0.1359	0.0088	0.009

Table A.7

REGRESSION DIAGNOSTICS FOR TOA EQUATION

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB>F
MODEL	3	39482.257	13160.752	179.243	0.0001
ERROR	25	1835.605	73.424220		
C TOTAL	28	41317.862			

ROOT MSE	8.568793	R-SQUARE	0.9556
DEP MEAN	62.931034	ADJ R-SQ	0.9502
C.V.	13.61616		

VARIABLE	DF	PARAMETER ESTIMATE	STANDARD ERROR	T FOR H0: PARAMETER=0	PROB > ABS(T)
INTERCEP	1	-46.917999	18.003362	-2.606	0.0152
THRWGT	1	8.727609	1.403500	6.218	0.0001
TEMP	1	0.046289	0.00730724	6.335	0.0001
SFCMIL	1	-31.263928	8.037125	-3.890	0.0007

COLLINEARITY DIAGNOSTICS VARIANCE PROPORTIONS

NUMBER	EIGENVALUE	CONDITION INDEX	PORTION INTERCEP	PORTION THRWGT	PORTION TEMP	PORTION SFCMIL
1	3.772	1.000	0.0005	0.0041	0.0006	0.0028
2	0.198962	4.354	0.0012	0.1700	0.0008	0.0971
3	0.024266	12.468	0.0569	0.6399	0.1164	0.5144
4	0.004405	29.264	0.9414	0.1860	0.8822	0.3856

Table A.8

REGRESSION DIAGNOSTICS FOR TOA EQUATION

ID	RSTUDENT	HAT DIAG H	COV RATIO	DFFITS	DFBETAS INTERCEP	DFBETAS THRWGT	DFBETAS TEMP	DFBETAS SFCMIL	COOK'S D
F100	-1.8891	0.2366	0.8858	-1.0517	0.6000	-0.3256	-0.4311	-0.3543	0.251
F101	-0.3661	0.2109	1.4591	-0.1893	0.0960	0.0149	-0.1158	-0.0068	0.009
F404	3.0468	0.2111	0.4035	1.5762	-0.9712	0.5227	0.6393	0.7047	0.467
TF30	0.5020	0.0772	1.2236	0.1452	0.0520	-0.0171	0.0001	-0.0941	0.005
TF33	0.1012	0.2282	1.5229	0.0550	0.0489	0.0098	-0.0347	-0.0462	0.001
TF34	-0.2509	0.1999	1.4562	-0.1254	-0.0534	-0.0207	0.0212	0.0890	0.004
TF39	-2.0590	0.2563	0.8259	-1.2087	-0.1390	0.4524	-0.4013	0.7921	0.323
TF41	1.6805	0.1113	0.8490	0.5948	-0.1029	-0.3140	0.3465	-0.2045	0.082
J30	-0.4867	0.0822	1.2333	-0.1456	-0.0252	0.0079	0.0345	-0.0416	0.005
J31	-1.0896	0.1116	1.0926	-0.3863	0.1065	0.1420	-0.0981	-0.2007	0.037
J33	-0.5421	0.1004	1.2465	-0.1811	0.0494	0.0681	-0.0480	-0.0906	0.008
J34	-0.5101	0.0626	1.2030	-0.1318	-0.0513	-0.0001	0.0495	-0.0052	0.004
J35	-0.7983	0.1005	1.1787	-0.2669	0.0273	0.1856	-0.0975	-0.0274	0.018
J40	0.8625	0.0535	1.1010	0.2051	0.0277	0.0067	-0.0423	0.0568	0.011
J42	0.1556	0.1109	1.3189	0.0549	0.0029	0.0168	-0.0189	0.0287	0.001
J46	0.2391	0.0505	1.2283	0.0552	0.0219	0.0082	-0.0232	0.0013	0.001
J47	-0.6249	0.0979	1.2236	-0.2058	0.0489	0.1377	-0.0925	-0.0473	0.011
J48	-0.6126	0.0631	1.1811	-0.1590	0.0271	0.0030	-0.0046	-0.0846	0.006
J52	1.9147	0.0734	0.7196	0.5390	0.4030	0.2016	-0.3671	-0.2505	0.066
J57	-0.4128	0.1040	1.2774	-0.1407	-0.0678	0.0817	0.0004	0.0856	0.005
J58	0.2345	0.3874	1.9043	0.1865	-0.1708	-0.0248	0.1358	0.1500	0.009
J60	-0.6006	0.3063	1.5989	-0.3991	-0.1342	-0.3708	0.2883	-0.0546	0.041
J65	0.6353	0.0641	1.1768	0.1662	0.0635	-0.0812	-0.0045	-0.0574	0.007
J69	1.1477	0.0601	1.0116	0.2902	-0.0064	-0.0274	-0.0123	0.1123	0.021
J71	0.1234	0.0965	1.2998	0.0403	0.0033	-0.0313	0.0151	-0.0123	0.000
J73	1.0941	0.0766	1.0495	0.3151	0.0767	-0.2036	0.0503	-0.0978	0.025
J75	0.1312	0.0732	1.2668	0.0369	0.0286	0.0090	-0.0231	-0.0205	0.000
J79	-1.1443	0.0579	1.0105	-0.2836	-0.1342	-0.1601	0.1635	0.0423	0.020
J85	-0.4275	0.3358	1.7195	-0.3039	-0.0609	-0.2840	0.1937	-0.0837	0.024